

北京大学物理学丛书
The Series of Advanced Physics of Peking University

近代光学信息处理

北京大学出版社 PEKING UNIVERSITY PRESS

北京大学物理学丛书

近代光学信息处理

宋菲君
S. Jutamulia 编著



北京大学出版社
北京

The Series of Advanced Physics of Peking University

Advanced Optical Information Processing

Song Feijun
S. Jutamulia

Peking University Press
Beijing

图书在版编目(CIP)数据

近代光学信息处理/宋菲君编著. --北京:北京大学出版社,1998.4

(北京大学物理学丛书)

ISBN 7-301-03700-7

Ⅰ.近… Ⅱ.宋… Ⅲ.近代光学-信息处理 Ⅳ.O438

中国版本图书馆 CIP 数据核字(98)第 07111 号

书　　名:近代光学信息处理

著作责任者:宋菲君　S. Jutamulia　编著

责 任 编 辑:瞿　定

标 准 书 号:ISBN 7-301-03700-7/O・414

出　版　者:北京大学出版社

地　　址:北京市海淀区中关村北京大学校内　100871

电　　话:出版部 62752015　发行部 62754140　编辑部　62752032

排　印　者:北京大学印刷厂

发　行　者:北京大学出版社

经　销　者:新华书店

850×1168　32 开本　11 印张　28.4 千字

1998 年 4 月第一版　1998 年 4 月第一次印刷

印　　数:0001-2 000 册

定　　价:18.50 元

内 容 提 要

本书论述光学信息处理的各个分支和相关领域的基本理论及重要应用，并介绍这一学科的前沿领域的研究成果和近期发展动向.

全书共分九章，内容包括傅里叶光学基础、经典光学信息处理、非相干光学信息处理、光学特征识别、广义(分数阶)傅里叶变换及其光学实现、光学小波变换、空间光调制器、光学神经网络和光折变介质信息处理.

本书可作为各类大学物理系、光学工程系、信息科学系高年级学生和研究生的教学参考书，也可供有关专业从事研究、开发的科技人员参考.

《北京大学物理学丛书》
编委会名单

前　言

物理学是自然科学的基础，是探讨物质结构和运动基本规律的前沿学科．几十年来，在生产技术发展的要求和推动下，人们对物理现象和物理学规律的探索研究不断取得新的突破．物理学的各分支学科有着突飞猛进的发展，丰富了人们对物质世界物理运动基本规律的认识和掌握，促进了许多和物理学紧密相关的交叉学科和技术学科的进步．物理学的发展是许多新兴学科、交叉学科和新技术学科产生、成长和发展的基础和前导．

为适应现代化建设的需要，为推动国内物理学的研究、提高物理教学水平，我们决定推出《北京大学物理学丛书》，请在物理学前沿进行科学研究和教学工作的著名物理学家和教授对现代物理学各分支领域的前沿发展做系统、全面的介绍，为广大物理学工作者和物理系的学生进一步开展物理学各分支领域的探索研究和学习，开展与物理学紧密相关的交叉学科和技术学科的研究和学习提供研究参考书、教学参考书和教材．

本丛书分两个层次．第一个层次是物理系本科生的基础课教材，这一教材系列，将在几十年来几代教师，特别是在北京大学教师的教学实践和教学经验积累的基础上，力求深入浅出、删繁就简，以适于全国大多数院校的物理系使用．它既吸收以往经典的物理教材的精华，尽可

能系统地、完整地、准确地讲解有关的物理学基本知识、基本概念、基本规律、基本方法；同时又注入科技发展的新观点和方法，介绍物理学的现代发展，使学生不仅能掌握物理学的基础知识，还能了解本学科的前沿课题和研究动向，提高学生的科学素质. 第二个层次是研究生教材、研究生教学参考书和专题学术著作. 这一系列将集中于一些发展迅速、已有开拓性进展、国际上活跃的学科方向和专题，介绍该学科方向的基本内容，力求充分反映该学科方向国内外前沿最新进展和研究成果. 学术专著首先着眼于物理学的各分支学科，然后再扩展到与物理学紧密相关的交叉学科.

愿这套丛书的出版既能使国内著名物理学家和教授有机会将他们的累累硕果奉献给广大读者，又能对物理的教学和科学研究起到促进和推动作用.

《北京大学物理学丛书》编辑委员会

1997 年 3 月

Preface

Physics is the foundation of natural sciences, a leading discipline of studying structures of matter and basic laws of motion. For several decades, driving by the demands of developing technology, the breakthrough in the studies of physical phenomenon and the laws of physics never end. During this period, all branches of Physics grew very fast and our knowledge of the basic laws governing the motion of the physical world was highly enriched. The growing of physics accelerated the progress of many physics related areas and technologies. The development of physics provided grounds and guidance for the birth and the growth of those new branches of physics, related areas and new technologies.

In order to catch up the main stream of the modernization and to give an impetus to scientific research and to improve teaching of physics in China. We decided to publish "*The Series of Advanced Physics of Peking University*". We invited those distinguished physicists and professors who worked in the frontier of physics to give series introductions to all branches of modern physics and recent developments in these fields. This series, as a consequence, provides textbooks and references for physicists and physics students in their studies of all branches of physics, related areas and technologies.

This series is divided into two sub-series of different levels, the first sub-series includes the textbooks of undergraduate

physics written by experienced teachers in Peking University in past decades. These textbooks were written concisely with deep insights and easier expressions, which adopt essences of physics textbook classics, explain fundamental concepts, laws and methods of physics in a systematic and rigorous way. In addition, these textbooks properly introduced the new approaches and the latest developments of physics for educational purposes. This sub-series is suitable for teaching of undergraduate physics for most universities and institutes in China. The second sub-series includes graduate textbook, references and academic writings. This sub series focuses on the latest developments and accomplishments in the active subjects of relevant research with international interests and introductions to those of fast developing research fields. The topics of academic writings mainly cover all branches of physics, but it will be generalized to closely related areas.

We wish the publication of this series could provide an opportunity for leading physicists and physics professors in China to show their fruitful accomplishments to general audience and to give an impetus to teaching and research in physics.

Editorial board of
"The Series of Advanced Physics of Peking University"
March 1997

序

世纪之交的物理学

20 世纪即将过去.物理学的革命,这场革命推动的整个自然科学和应用技术的伟大变革,以及这些变革对人类社会的影响,将作为这个世纪的一个重要标志而载入史册.这段令人神往的历史,给正处在世纪之交的我们以什么样的启示呢?

首先的启示是作为研究物质结构和运动的基本规律的物理学,总是生机勃勃、不断地开辟自己前进的道路的.1803 年道尔顿提出了近代的原子论,认为世间万物都是由几十种不同种类的原子(那时只认识到三十来种)组成的.经过近一个世纪多方面的研究和争论,科学界接受了和发展了这个学说.到 19 世纪 60 年代,元素的数目增到六十多种,而且还认识到不同元素的性质是有内在联系的,门捷列夫的周期律描述了这种联系.19 世纪末,物理学家们发现了电子、α 粒子、放射性、X 射线……发现原子是可以改变的,原子不是物质构成的最小单元.20 世纪初卢瑟福建立了原子结构的“行星”模型.探讨原子结构模型和经典物理学之间的矛盾,导致了量子力学的诞生,产生了现代原子、分子物理、凝聚态物理、原子核物理……历史也许有某种类似性.在这个世纪之交,物理学又

正在进入一个新的层次.本世纪50年代,人们找到作为构成物质的基石的基本粒子有三十来种.也认识到它们之间的相互作用和相互转化.探索这些“基本”粒子的更深层次的构造的努力,近三十年来取得十分辉煌的成就.三代夸克和三代轻子的粒子模型、电弱统一理论和量子色动力学,这被称作粒子物理的标准模型的建立以及它在各方面的成功,正是标志着物理学正在进入物质世界的一个更深的层次,毫无疑问这将是物理学历史上一个具有划时代意义的大事.但是,大多数的物理学家都会同意,从本质上说,目前的标准模型还是一个唯象的模型.在欢呼它取得的多方面的胜利时,也要看到同时提出了一系列带本质性的问题.也许可以说,进入这个新层次将带来的最本质的新的物理,还没有来到我们中间.19世纪末的物理学家没能猜测到,进入比原子更深层次的探索,会在什么时候和在哪点上带来新的物理.20世纪末的人们也不能奢望会比前辈高明多少!

20世纪物理学的革命,表现出人类理性思维的伟大胜利.狭义相对论,特别是广义相对论,以它深邃的思考,严整的形式和美丽的表述,震撼着一代又一代的物理学工作者的心灵.但是,爱因斯坦也在他那无与伦比的思考导致的宇宙模型面前困惑了.感谢近四十多年来一大批物理学家和天文学家辛勤的努力和非凡的勇气,这个难以想象的革命性的关于宇宙的概念和图象建立起来了,并得到科学界多数人的认同.在大致一百五十多亿年前,宇宙从一个具有无限大的密度和具有无限大的时空曲率的点开始了.人们猜想,在宇宙膨胀,密度和温度降低中,

到10^{-44}秒时，重力相互作用和其他相互作用分离开来；到10^{-36}秒时，强相互作用和电弱相互作用分离；直到10^{-10}秒时，弱相互作用才与电磁相互作用分离，世界变成了我们现在所处的有四种基本相互作用的世界. 到10^{-5}秒时，夸克开始结合成强子，也许应当说，只有从此之后的物理才是当代物理学家可以比较有把握来谈论的. 直到10^{12}秒（也就是三万多年）后，原子才开始出现. 这样一个综合了亚核子尺度（小于10^{-14}厘米）的物理和宇观尺度（大于10^{24}厘米）的物理的宇宙演化模型的建立，可以说是人类认识史上一个最具有革命性的、划时代的伟大事件，当代人还难以全部理解它的意义. 这个宇宙学的标准模型和近年来天体物理学家取得的辉煌的成就，在物理学的面前提出了十分严峻的具有本质的挑战. 如何理解这样有限而无界的时空和它的奇点？ 什么是在这样演化中的物质和运动规律？为什么宇宙学中有那么多"巧合"？……总之，面对着如此壮观而又如此神奇的宇宙之迹，除了由衷的赞美和敬佩，人们不禁想起爱因斯坦的话："自然界最不可理解的就是它竟然可以理解！"

本世纪物理学的发展给我们的又一个教益是：物质世界是有层次的，反映物质世界的物理学规律也是有层次的. 每一层次的物理都植根于更深层次的物理学. 但是，每个层次的物理都是在真实的意义上不可穷尽的. 在大自然千姿百态的丰富性面前，那些断言某某学科将不会有什么发展的说法总是被事实所粉碎的. 经典力学、经典电动力学并不因为量子力学、量子电动力学的发展而被排斥出物理学，近年来我们还不停地学习它们新的、有

深刻意义的进展.光学和凝聚态物理学半个世纪来的巨大的、令人应接不暇的发展提供了最能说服人的例子.也许人们常常由于赞颂它们在实际应用上的威力的同时,不够强调它们在物理学学科上的意义.但如果我们想到在当代粒子物理和宇宙学中最重要的观念,如相变、对称性破缺、拓扑性缺陷、红外发散等等,都首先来自凝聚态物理学,而当代凝聚态物理又都广泛地而且本质地使用量子场论的语言和技术时,我们就会确信物理学的丰富性、多样性和统一性,"只有一个物理学"!

在20世纪,物理学的基本概念和技术已被应用到所有的自然科学领域.物理学与其他自然科学学科之间的边缘领域,一定意义上是当代自然科学中最富于获得丰硕成果的机遇的领域.边缘领域的发展,又反过来丰富、加深和支持了物理学本身的发展.量子力学和现代物理实验技术的应用,大大推进了现代化学的发展.对分子结构、性能和反应机理的研究,又丰富和推动了现代物理的进步.而且,如果没有现代化学的支撑,现代物理学中好些分支学科都不能产生和发展.地球科学、生命科学与物理学的边缘领域的发展,也将会是类似的情景.特别令人感到兴奋的一个新事物是近二三十年复杂性科学的发展.数学、物理学,特别是物理学与化学、地球科学、生命科学、各种应用技术科学的边缘领域研究的发展,都使人们相信,在复杂性(多维度,多组元,非线性,非平衡和开放的)系统的结构、性能和演化中,有一些具有普遍性的运动规律和运动模式.人们自然回想到,开始于上一世纪中叶的研究能量守恒和转化的热力学和分子运动论的发

展，本世纪统计物理和涨落理论的发展.有理由相信一门有重要的基础科学意义的学科，复杂性物理正在形成.它现在已经显出可能对物理学中一些最基础的问题，如必然性和随机性、无序化的倾向和有序结构的生成，不同层次的结构的自相似性等，作出有深刻物理意义的回答.也许，历史会表明，这也是人类认识史上又一个划时代的事物，同时，无疑地会对化学、地球科学、生命科学、认知科学和各种应用技术发生巨大的影响.

物理学作为一门最基础的自然科学，它的发展动力是深深地植根于人类对真理的非功利的追求.但是，历史的发展将越来越有力地证明，正是这种非功利的追求给人类带来最大的收益.本世纪发生的主要源于物理学的进展的技术革命，就是最有说服力的例子.当代技术进步的主要推动力来自纯学科性的基础研究.研究室和实验室中纯学科性的研究转变为重要的应用技术，实际生产和社会发展中遇到的问题转化为有基础学科意义的研究课题，两者关系愈来愈密切，周期愈来愈短.与之相应，在现代，杰出的基础科学研究人材和优秀的应用技术开发人材在科学素质上的要求变得更加一致了.在世纪之交，无论是制造业还是服务业，也无论是材料、信息、能源、交通、环境等技术部门，都在呼唤着新的技术变革.而认真考察就会发现，多数这些变革都主要基于物理学近年的进展.21世纪物理学毫无疑问仍是技术进步的主要源泉.

物理学的发展从来就对人类社会思想、文化发生巨大影响.20世纪的物理学革命就更是这样.人类社会进步的一个主要动力便是科学精神，现代科学精神的典范和

集中的反映就是现代物理学.我国是一个文明古国,在历史上曾经对人类的文化和科学发展做出过光辉的贡献.但是,我国接受近代科学的时间还很短,现代科学的精神实质和思维方式扎根我国还要做艰苦的努力.有些人跟着西方一些比较浅薄的哲学流派的后面,宣扬一些贬低和反对现代科学精神的言论,甚至把当代社会中由于社会矛盾而造成的后果,都归罪到现代科学精神上.这当然是完全错误的.以现代物理学为代表的科学精神,是人类进步的一面旗帜,它将高高飘扬在未来的岁月中.

当我们站在新世纪的门槛上,回顾20世纪物理学的辉煌时,会更加确信,在21世纪物理学将会同样辉煌.那些有幸进入物理学工作者的行列的年青朋友,历史将会证明,你们的选择是完全正确的.

中国科学院院士 甘子钊

1997年4月

作者前言

光学信息处理是近年来发展起来的一门新兴学科. 1948 年全息术的发明, 1955 年光学传递函数概念的建立, 以及 1960 年新型的强相干光源——激光的诞生, 是近代光学发展史上的三件大事, 也是光学信息处理的基础, 而透镜的傅里叶变换效应则构成了光学信息处理的理论框架.

与其他形态的信号处理相比, 光学信息处理具有高度并行、大容量的特点. 近年来, 这一学科发展很快, 理论体系已日趋成熟、完整, 成为信息科学的重要分支, 在一些领域中逐渐进入实用阶段.

本书共分九章. 第一章"傅里叶光学基础", 用较少的篇幅讲授光学信息处理的理论基础; 第二章回顾了早期的研究工作, 这些工作成为近代光学信息处理的初步模型和预备概念. 从第三章起, 每一章讲述一个专题. 第三章讨论非相干光学信息处理的各种问题和应用; 第四章介绍一个重要的分支——光学特征识别. 五、六两章介绍本学科的两个前沿领域: 广义(分数阶)傅里叶变换的光学实现和光学小波变换, 它们正是目前信息科学最活跃的领域. 第七章介绍光学信息处理系统的关键器件——各种空间光调制器的原理、结构, 及其在非相干光信息处理——大屏幕信息显示中的重要应用. 第八章简要论述了光学信息处理的另一分支——光学神经网络. 第九章则讨论了光折变效应及相关的信息处理课题.

事实上, 几乎每一章都涉及这一学科的一个分支或一个领域, 在这些领域中已发表了大量的论文和一些专著. 作为近代物理学的丛书之一, 本书的宗旨是力求在有限的篇幅内, 讲述基本原理和有关的效应, 建立清晰的物理图象, 讨论数学处理方法, 介绍典型的应用. 对这些领域有兴趣的读者, 可以进一步阅读每章所附的参考文献.

宋菲君撰写了第一、五、六、七、九章，S. Jutamulia 撰写了第二、三、四、八章，由宋菲君译成中文并统稿.

本书由中国科学院物理研究所副所长聂玉昕研究员审查，中国大恒公司俞甯高级工程师对全书进行了校勘. 物理所朱镛研究员、叶佩弦研究员、张洪钧研究员提供了有价值的资料，并与作者进行了非常有益的讨论. 北京大学物理系吴崇试教授编写了部分附录，王威礼教授改写了第九章第二节. 王瑜工程师、杨颜峰博士、吕建伟先生、李勇先生和张斌小姐也参与了部分工作，作者在此表示由衷的感谢. 最后，作者还要感谢北京大学出版社对本书的出版给予的热情支持.

宋菲君，S. Jutamulia

1997,10

目　录

第一章　傅里叶光学基础 ………………………………………… (1)

1.1　二维傅里叶分析 ………………………………………… (1)

1.2　空间带宽积和测不准关系式 ………………………………… (9)

1.3　平面波的角谱和角谱的衍射 ………………………………… (15)

1.4　透镜系统的傅里叶变换性质 ………………………………… (20)

参考文献 ………………………………………………………… (25)

第二章　经典光学信息处理 ………………………………………… (26)

2.1　引言 ………………………………………………………… (26)

2.2　早期发展 …………………………………………………… (26)

2.3　傅里叶处理器 ……………………………………………… (29)

2.4　线性系统与卷积 …………………………………………… (31)

2.5　空间滤波 …………………………………………………… (33)

2.6　照相图像的恢复 …………………………………………… (34)

2.7　全息术 ……………………………………………………… (36)

2.8　傅里叶变换全息图 ………………………………………… (38)

2.9　相关和卷积 ………………………………………………… (41)

2.10　结论 ……………………………………………………… (43)

参考文献 ………………………………………………………… (44)

第三章　非相干光学信息处理 ……………………………………… (46)

3.1　杨氏干涉仪和空间相干性 ………………………………… (46)

3.2　非相干像的形成 …………………………………………… (49)

3.3　MTF 的测量 ……………………………………………… (50)

3.4　非相干空间滤波 …………………………………………… (52)

3.5　迈克耳孙干涉仪和时间相干性 …………………………… (54)

3.6　傅里叶变换光谱仪 ………………………………………… (55)

3.7　投影显示的消像素技术 …………………………………… (57)

3.8 计算层析技术 …………………………………………………… (59)
3.9 结论 …………………………………………………… (62)
参考文献 …………………………………………………… (62)
第四章 光学图像识别 …………………………………………………… (64)
4.1 图像识别和光学相关器 …………………………………………………… (64)
4.2 非相干识别器 …………………………………………………… (65)
4.3 Vander Lugt 相关器 …………………………………………………… (66)
4.4 实时 Vander Lugt 相关器 …………………………………………………… (69)
4.5 Vander Lugt 相关器的小型化 …………………………………………………… (72)
4.6 旋转不变 Vander Lugt 相关器 …………………………………………………… (75)
4.7 比例不变 Vander Lugt 相关器 …………………………………………………… (77)
4.8 联合变换相关器 …………………………………………………… (78)
4.9 实时联合变换相关器 …………………………………………………… (81)
4.10 联合变换相关器的应用 …………………………………………………… (84)
4.11 旋转不变联合变换相关器 …………………………………………………… (88)
4.12 结论 …………………………………………………… (91)
参考文献 …………………………………………………… (91)
第五章 广义傅里叶变换及其光学实现 …………………………………………………… (94)
5.1 引言 …………………………………………………… (94)
5.2 广义傅里叶变换的定义及性质 …………………………………………………… (96)
5.3 广义傅里叶变换的本征函数 …………………………………………………… (104)
5.4 用透镜系统实现广义傅里叶变换的基本光学单元 …………………………………………………… (106)
5.5 基本光学单元的组合 …………………………………………………… (113)
5.6 用自聚焦效应光波导实现广义傅里叶变换 …………………………………………………… (118)
5.7 维格纳变换 …………………………………………………… (125)
参考文献 …………………………………………………… (134)
第六章 光学小波变换 …………………………………………………… (136)
6.1 引言 …………………………………………………… (136)
6.2 从短时傅里叶变换到小波变换 …………………………………………………… (138)
6.3 小波变换的定义和性质 …………………………………………………… (144)

6.4　实现一维小波变换的光学系统 …………………………… (151)
6.5　用多通道匹配滤波实现二维小波变换 ……………………… (153)
6.6　光学小波变换匹配滤波器在图像识别中的应用 …………… (159)
6.7　光学 Haar 小波变换和图形边缘探测 ……………………… (165)
参考文献 ………………………………………………………… (172)
第七章　空间光调制器 ……………………………………… (174)
7.1　概论 …………………………………………………………… (174)
7.2　磁光空间光调制器(MOSLM) ……………………………… (180)
7.3　液晶的扭曲效应及薄膜晶体管驱动液晶显示器(TFT-LCD)
…………………………………………………………………… (182)
7.4　液晶显示器在非相干光信息处理中的应用
大屏幕投影电视 ……………………………………………… (190)
7.5　液晶光阀 ……………………………………………………… (192)
7.6　线性电光效应和 PROM 器件 ……………………………… (198)
7.7　数字微反射镜器件(DMD)和数字化投影 ………………… (208)
参考文献 ………………………………………………………… (215)
附录 A7　偏振光在线性扭曲介质中的传播 ……………………… (218)
第八章　光学神经网络 ……………………………………… (223)
8.1　引言 …………………………………………………………… (223)
8.2　一般基础 ……………………………………………………… (224)
8.3　光学互连 ……………………………………………………… (228)
8.4　神经网络的学习 ……………………………………………… (242)
8.5　结论 …………………………………………………………… (252)
参考文献 ………………………………………………………… (253)
第九章　光折变介质信息处理 ……………………………… (259)
9.1　相位共轭和非线性介质中的相位共轭效应 ………………… (259)
9.2　光折变效应相位共轭器 ……………………………………… (265)
9.3　用耦合波近似处理光折变晶体中的二波混频和四波混频
效应 …………………………………………………………… (273)
9.4　自泵浦相位共轭效应 ………………………………………… (283)

9.5 互泵浦相位共轭及双色泵浦光折变振荡……………………（294）
9.6 利用图像对相位光栅的编码实现非相干光-相干光转换
……………………………………………………………………（305）
9.7 相位共轭在联合傅里叶变换相关识别中的应用……………（308）
9.8 光折变非线性联合变换相关器……………………………（311）
参考文献 ……………………………………………………………（315）
附录 A9 耦合波方程的解 ………………………………………（320）

Contents

1 Elements of the Fourier optics .... (1)

1.1 Two-dimensional Fourier analysis .... (1)

1.2 Space-bandwidth product and uncertainty principle .... (9)

1.3 Angular spectrum of plane wave and its diffraction .... (15)

1.4 Fourier transform properties of lens systems .... (20)

References .... (25)

2 Classical information processing .... (26)

2.1 Introduction .... (26)

2.2 Early development .... (26)

2.3 Fourier processor .... (29)

2.4 Linear system and convolution .... (31)

2.5 Spatial filtering .... (33)

2.6 Restoration of photograph .... (34)

2.7 Holography .... (36)

2.8 Fourier transform hologram .... (38)

2.9 Correlation and convolution .... (41)

2.10 Concluding remarks .... (43)

References .... (44)

3 Incoherent optical processing .... (46)

3.1 Young interferometer and spatial coherence .... (46)

3.2 Incoherent image formation .... (49)

3.3 Measurement of MTF .... (50)

3.4 Incoherent spatial filtering .... (52)

3.5 Michelson interferometer and temporal coherence .... (54)

3.6 Fourier transform spectroscopy .... (55)

3.7 Depixelization of projection display .... (57)

3.8 Computed tomography ………… (59)
3.9 Concluding remarks ………… (62)
References ………… (62)
4 Optical pattern recognition ………… (64)
4.1 Pattern recognition and optical correlator ………… (64)
4.2 Incoherent correlator ………… (65)
4.3 Vander Lugt correlator ………… (66)
4.4 Real-time Vander Lugt correlators ………… (69)
4.5 Compact Vander Lugt correlators ………… (72)
4.6 Rotation invariant Vander Lugt correlators ………… (75)
4.7 Scale invariant Vander Lugt correlator ………… (77)
4.8 Joint transform correlator ………… (78)
4.9 Real-time joint transform correlators ………… (81)
4.10 Applications of joint transform correlators ………… (84)
4.11 Rotational invariant joint transform correlators ………… (88)
4.12 Concluding remarks ………… (91)
References ………… (91)
5 Fractional Fourier transform and its optical implementation ………… (94)
5.1 Introduction ………… (94)
5.2 Definition and properties of fractional Fourier transform ………… (96)
5.3 Eigenfunctions of fractional Fourier transform ………… (104)
5.4 Implementation of fractional Fourier transform using lens systems ………… (106)
5.5 Combination of basic optical units ………… (113)
5.6 Implementation of fractional Fourier transform using self-focusing medium ………… (118)
5.7 Wigner transform ………… (125)
References ………… (134)

6 Optical wavelet transform ······ (136)
6. 1 Introduction ······ (136)
6. 2 Short temporal Fourier transform and wavelet transform ······ (138)
6. 3 Definition and properties of wavelet transform ······ (144)
6. 4 Optical implementation of one-dimensional wavelet transform ······ (151)
6. 5 Implementation of two-dimensional wavelet transform using multi-channel matched filtering ······ (153)
6. 6 Applications of wavelet transform based matched filtering to pattern recognition ······ (159)
6. 7 Edge detection by optical Haar wavelet transform ······ (165)
References ······ (172)
7 Spatial light modulators ······ (174)
7. 1 Introduction ······ (174)
7. 2 Magneto-optical spatial light modulators ······ (180)
7. 3 Twisted liquid crystals and liquid crystal display ······ (182)
7. 4 Liquid crystal projection display ······ (190)
7. 5 Liquid crystal light valve ······ (192)
7. 6 Pockels readout optical modulator ······ (198)
7. 7 Digital micromirror device ······ (208)
References ······ (215)
Appendix: Propagation of light wave in twisted medium ······ (218)
8 Optical neural networks ······ (223)
8. 1 Introduction ······ (223)
8. 2 General background ······ (224)
8. 3 Optical implementation ······ (228)
8. 4 Learning in neural networks ······ (242)
8. 5 Concluding remarks ······ (252)
References ······ (253)

9 Optical information processing in the photorefractive crystals ······ (259)

9.1 Phase conjugation in nonlinear medium ······ (259)

9.2 Phase conjugation in photorefractive medium ······ (265)

9.3 Two-wave mixing and four-wave mixing coupling ······ (273)

9.4 Self-pumped phase conjugation ······ (283)

9.5 Mutually pumped phase conjugation and double-color pumped oscillation ······ (294)

9.6 Incoherent-to-coherent conversion ······ (305)

9.7 Applications of phase conjugation to joint transform correlation ······ (308)

9.8 Phtorefractive nonlinear joint transform correlator ······ (311)

References ······ (315)

Appendix: Solution to coupling wave equations ······ (320)

第一章　傅里叶光学基础

1.1　二维傅里叶分析

1.1.1　定义及存在条件

一个复变函数 $g(x,y)$ 的傅里叶变换可表为

$$G(u,v)=\mathscr{F}\{g(x,y)\}$$
$$=\iint_{-\infty}^{\infty} g(x,y)\exp[-\mathrm{i}2\pi(ux+vy)]\mathrm{d}x\mathrm{d}y, \tag{1}$$

通常称 $g(x,y)$ 为原函数，$G(u,v)$ 为变换函数或像函数，它一般也是复变函数．(1)式的逆变换为

$$g(x,y)=\mathscr{F}^{-1}\{G(u,v)\}$$
$$=\iint_{-\infty}^{\infty} G(u,v)\exp[\mathrm{i}2\pi(ux+vy)]\mathrm{d}u\mathrm{d}v. \tag{2}$$

变换存在的条件为：

(1) $g(x,y)$ 在全平面绝对可积；

(2) $g(x,y)$ 在全平面只有有限个间断点，在任何有限的区域内只有有限个极值；

(3) $g(x,y)$ 没有无穷大型间断点．

以上条件并非必要，实际上，"物理的真实"就是变换存在的充分条件．

以下我们常用 $g(x,y)\Leftrightarrow G(u,v)$ 表示变换对．对于光学傅里叶变换，x,y 是空间变量，u,v 则是空间频率变量．在一维情况下，有时也用希腊字母 ν 表示频率变量．关于光学傅里叶变换，在本章 1.3 节、1.4 节及第二、第五章将详细介绍．

1.1.2 δ函数的傅里叶变换

由 δ 函数的定义容易得到

$$\delta(x - x_0, y - y_0) \Leftrightarrow \exp[-i2\pi(ux_0 + vy_0)], \tag{3}$$

当 $x_0=0, y_0=0$ 时得到

$$\delta(x, y) \Leftrightarrow 1. \tag{4}$$

上式有深刻的物理意义，它表示点源函数具有权重为 1 的最丰富的频谱分量. 因此光学中常用点光源来检测系统的响应特性，即脉冲响应. (3)式还可表为

$$\delta(x - x_0, y - y_0) = \iint_{-\infty}^{\infty} \exp\{-i2\pi[u(x - x_0) + v(y - y_0)]\} du dv, \tag{5}$$

它正是 δ 函数的积分表达式.

根据 δ 函数的偏导数的定义

$$\int_{-\infty}^{\infty} \delta^{(n)}(x) g(x) dx = (-1)^n g^{(n)}(0), \tag{6}$$

得到 $\delta^{(k,l)}(x,y)$ 的傅里叶变换：

$$\delta^{(k,l)}(x,y) = \frac{\partial^{k+l}}{\partial x^k \partial y^l} \delta(x,y) \Leftrightarrow (i2\pi u)^k (i2\pi v)^l. \tag{7}$$

1.1.3 傅里叶变换的基本性质

以下设 A, B 为任意常数，a, b 为非零实数，k, l, n 为自然数，G, H 分别为 g, h 的傅里叶变换.

傅里叶变换的基本性质如下：

(1) 线性(linearity)

$$Ag(x,y) + Bh(x,y) \Leftrightarrow AG(u,v) + BH(u,v). \tag{8}$$

(2) 缩放及反演(scaling and inversion)

$$g(ax, by) \Leftrightarrow \frac{1}{|ab|} G\left(\frac{u}{a}, \frac{v}{b}\right). \tag{9}$$

上式表明空域信号的展宽将引起频域信号的压缩. 特别是当 $a=b=-1$ 时,得到反演的变换性质:

$$g(-x,-y) \Leftrightarrow G(-u,-v). \tag{10}$$

(3) 位移 (shift)

$$g(x+x_0, y+y_0) \Leftrightarrow \exp[\mathrm{i}2\pi(ux_0+vy_0)]G(u,v). \tag{11}$$

表示原函数的位移引起变换函数的相移.

(4) 共轭 (conjugation)

$$g^*(x,y) \Leftrightarrow G^*(-u,-v). \tag{12}$$

(5) 卷积 (convolution)

函数 $g(x,y)$ 和 $h(x,y)$ 的卷积定义如下:

$$g(x,y) * h(x,y) = \iint_{-\infty}^{\infty} g(\xi,\eta)h(x-\xi, y-\eta)\mathrm{d}\xi\mathrm{d}\eta. \tag{13}$$

容易证明两个函数卷积的变换等于它们的变换函数的积,即

$$g(x,y) * h(x,y) \Leftrightarrow G(u,v)H(u,v). \tag{14}$$

δ 函数的卷积有特殊的性质:

$$g(x) * \delta(x-x_0) = g(x-x_0). \tag{15}$$

$$g(x,y) * \delta^{(k,l)}(x,y) = g^{(k,l)}(x,y). \tag{16}$$

(6) 导数的变换公式可由(7)式导出

$$g^{(k,l)}(x,y) \Leftrightarrow (\mathrm{i}2\pi u)^k(\mathrm{i}2\pi v)^l G(u,v). \tag{17}$$

(7) 相关(correlation)

函数 $g(x,y)$ 和 $h(x,y)$ 的相关定义为

$$g(x,y) \otimes h(x,y) = \iint_{-\infty}^{\infty} g^*(\xi,\eta)h(x+\xi, y+\eta)\mathrm{d}\xi\mathrm{d}\eta. \tag{18}$$

当 $g=h$ 时成为自相关,有

$$g(x,y) \otimes g(x,y) = \iint_{-\infty}^{\infty} g^*(\xi,\eta)g(x+\xi, y+\eta)\mathrm{d}\xi\mathrm{d}\eta, \tag{19}$$

相关的变换可以利用卷积的变换公式导出:

$$\begin{aligned} g(x,y) \otimes h(x,y) &= g^*(-x,-y) * h(x,y) \\ &\Leftrightarrow G^*(u,v)H(u,v), \end{aligned} \tag{20}$$

$$g(x,y) \otimes g(x,y) \Leftrightarrow |G(u,v)|^2. \qquad (21)$$

(8) 矩 (moment)

$g(x,y)$的(k,l)阶矩定义为

$$M_{k,l} = \iint_{-\infty}^{\infty} g(x,y) x^k y^l \mathrm{d}x\mathrm{d}y. \qquad (22)$$

将逆变换表达式(2)代入上式，得到

$$M_{k,l} = \iint_{-\infty}^{\infty} G(u,v)\mathrm{d}u\mathrm{d}v \iint_{-\infty}^{\infty} x^k y^l \exp[\mathrm{i}2\pi(ux+vy)]\mathrm{d}x\mathrm{d}y. \qquad (23a)$$

由 δ 函数导数的变换表达式(7)，上式内部的积分

$$\iint_{-\infty}^{\infty} x^k y^l \exp[\mathrm{i}2\pi(ux+vy)]\mathrm{d}x\mathrm{d}y = (\mathrm{i}2\pi)^{-k-l}\delta^{(k,l)}(u,v).$$

代入(23a)，再用一次 δ 函数导数的定义式(6)，得到矩的表达式

$$M_{k,l} = (-\mathrm{i}2\pi)^{-k-l} G^{(k,l)}(0,0). \qquad (23b)$$

(9) Parseval 定理

(21)式可用逆变换表达式改写为

$$\iint_{-\infty}^{\infty} g^*(\xi,\eta)h(x+\xi,y+\eta)\mathrm{d}\xi\mathrm{d}\eta$$

$$= \iint_{-\infty}^{\infty} G^*(u,v)H(u,v)\exp[\mathrm{i}2\pi(ux+vy)]\mathrm{d}u\mathrm{d}v, \qquad (24)$$

令 $x=y=0$，上式为

$$\iint_{-\infty}^{\infty} g^*(\xi,\eta)h(\xi,\eta)\mathrm{d}\xi\mathrm{d}\eta = \iint_{-\infty}^{\infty} G^*(u,v)H(u,v)\mathrm{d}u\mathrm{d}v, \qquad (25)$$

这一关系式称为 Parseval 定理. 当 $h=g$ 时，上式化为

$$\iint_{-\infty}^{\infty} |g(\xi,\eta)|^2\mathrm{d}\xi\mathrm{d}\eta = \iint_{-\infty}^{\infty} |G(u,v)|^2 \mathrm{d}u\mathrm{d}v, \qquad (26)$$

该式又称完备关系式，实际上是能量守衡定律在空域和频域中表达式一致性的表现.

1.1.4 特殊函数及其傅里叶变换

(1) $\mathrm{rect}(x)$，$\Lambda(x)$及 $\mathrm{sinc}(x)$

这三个函数定义如下：

$$\mathrm{rect}(x)=\begin{cases}1, & |x|\leqslant 1/2,\\ 0, & \text{其他},\end{cases} \tag{27}$$

$$\Lambda(x)=\begin{cases}1-|x|, & |x|\leqslant 1,\\ 0, & \text{其他},\end{cases} \tag{28}$$

$$\mathrm{sinc}(x)=\sin\pi x/\pi x. \tag{29}$$

它们的傅里叶变换分别为

$$\mathrm{rect}(x)\Leftrightarrow\mathrm{sinc}(u), \tag{30}$$

$$\mathrm{sinc}(x)\Leftrightarrow\mathrm{rect}(u), \tag{31}$$

$$\Lambda(x)\Leftrightarrow\mathrm{sinc}^2(u). \tag{32}$$

(2) $\mathrm{sgn}(x)$及 $\mathrm{step}(x)$

符号函数 $\mathrm{sgn}(x)$及阶跃函数 $\mathrm{step}(x)$定义为

$$\mathrm{sgn}(x)=\begin{cases}1, & x>0,\\ 0, & x=0,\\ -1, & x<0,\end{cases} \tag{33}$$

$$\mathrm{step}(x)=\begin{cases}1, & x\geqslant 0,\\ 0, & x<0.\end{cases} \tag{34}$$

$\mathrm{sgn}(x)$的傅里叶变换为

$$\mathrm{sgn}(x)\Leftrightarrow\frac{1}{\mathrm{i}\pi u}, \tag{35}$$

该结果是令 $g(x)=\mathrm{e}^{-\alpha x}(\alpha>0,x>0)$及$-\mathrm{e}^{\alpha x}(x<0)$，先求出 g 的变换表达式，再令 $\alpha\to 0$ 得到的. $\mathrm{step}(x)$的傅里叶变换为

$$\mathrm{step}(x)=\frac{1}{2}\mathrm{sgn}(x)+\frac{1}{2}\Leftrightarrow\frac{1}{\mathrm{i}2\pi u}+\frac{1}{2}\delta(u). \tag{36}$$

利用 $\mathrm{step}(x)$的变换式及卷积定理，可求出积分$\int_{-\infty}^{x}g(\xi)\mathrm{d}\xi$的变换：

$$\int_{-\infty}^{x} g(\xi)\mathrm{d}\xi = \int_{-\infty}^{\infty} g(\xi)\mathrm{step}(x-\xi)\mathrm{d}\xi$$
$$= g(x) * \mathrm{step}(x)$$
$$\Leftrightarrow G(u)\left[\frac{1}{\mathrm{i}2\pi u} + \frac{1}{2}\delta(u)\right]$$
$$= \frac{G(u)}{\mathrm{i}2\pi u} + \frac{G(0)\delta(u)}{2}. \qquad (37)$$

(3) 周期函数

设函数 $g(x)$可展开为傅里叶级数

$$g(x) = \sum_{n=-\infty}^{\infty} C_n \exp(\mathrm{i}2n\pi f_0 x), \qquad (38)$$

式中系数

$$C_n = \frac{1}{X}\int_{-X/2}^{X/2} g(x)\exp(-\mathrm{i}2n\pi f_0 x)\mathrm{d}x, \qquad (39)$$

周期 $X=1/f_0$. 对(38)式两边取傅里叶变换得到

$$G(u) = \sum_{n=-\infty}^{\infty} C_n \delta(u - nf_0), \qquad (40)$$

在推导中用到 δ 函数的积分变换式.

(4) $\mathrm{comb}(x)$

$\mathrm{comb}(x)$定义为

$$\mathrm{comb}(x) = \sum_{n=-\infty}^{\infty} \delta(x-n), \qquad (41)$$

为一系列间隔为 1 的 δ 函数组成. 它可以用周期函数的傅里叶级数表达为

$$\mathrm{comb}(x) = \sum_{n=-\infty}^{\infty} \exp(\mathrm{i}2n\pi x), \qquad (42)$$

系数 $C_n=1$. 因此由(40)式可得

$$\mathrm{comb}(x) \Leftrightarrow \mathrm{comb}(u). \qquad (43)$$

设 X 为实数常数, 则有

$$(1/X)g(x) * \mathrm{comb}(x/X) = \sum_{n=-\infty}^{\infty} g(x-nX), \qquad (44)$$

结果得到了以 $nX(n=0,\pm1,\pm2,\cdots)$ 为中心的一系列重复出现的波形 $g(x-nX)$，这一现象称为“复现”.

函数 $g(x)$ 与 $\mathrm{comb}(x/X)$ 的乘积用 $g_s(x)$ 表示，

$$\begin{aligned} g_s(x) &= g(x)\mathrm{comb}(x/X) \\ &= X\sum_{n=-\infty}^{\infty} g(nX)\delta(x-nX), \end{aligned} \tag{45}$$

g_s 称 g 的抽样函数，X 为抽样间隙，$x_n=nX$ 称样点，$g(x_n)$ 称样值. 所以 $g(x)$ 的抽样函数 $g_s(x)$ 是以样值为权重的 δ 函数序列.

1.1.5 功率谱与空间自相关函数

由 Parseval 定理(26)式

$$\iint_{-\infty}^{\infty} |g(x,y)|^2 \mathrm{d}x\mathrm{d}y = \iint_{-\infty}^{\infty} |G(u,v)|^2 \mathrm{d}u\mathrm{d}v,$$

如 $g(x,y)$ 为光场的复振幅分布，则 $|g(x,y)|^2$ 代表光强分布，$|G(u,v)|^2$ 则表示单位频率间隔的光能量，称为功率谱，用 $s(u,v)$ 表示为

$$s(u,v) = |G(u,v)|^2. \tag{46}$$

根据变换定理，我们得到

$$g(x,y) \otimes g(x,y) \Leftrightarrow |G(u,v)|^2 = s(u,v). \tag{47}$$

$g\otimes g$ 在光学上称为空间自相关函数. 上式表示功率谱是空间自相关函数的傅里叶变换. 空间自相关函数表征空间相距为 (x,y) 的两点之间场的相似性或关联性. 它是场的空间相干性的度量. 场的相干性较高时，功率谱 s 的弥散就较小，表示光功率在频域内集中在很小的区域中(这样的光波可称为准单色光)；反之当场的相干性较差时，功率谱 s 的弥散就较大，表示光功率在频域中分布在较大的区域内，包含较宽的波段.

1.1.6 二维傅里叶变换表

常用的二维傅里叶变换见表 1.1.

表 1.1　二维傅里叶变换

$g(x,y)$	$G(u,v)$
$g(-x,-y)$	$G(-u,-v)$
$g^{*}(x,y)$	$G^{*}(-u,-v)$
$G(x,y)$	$g(-u,-v)$
$G^{*}(\pm x,\pm y)$	$g^{*}(\mp u,\mp v)$
$g(ax,by)$	$\frac{1}{\lvert ab\rvert}G\left(\frac{u}{a},\frac{v}{b}\right)$
$g(x\pm x_0,y\pm y_0)$	$\exp[\pm i2\pi(ux_0+vy_0)]G(u,v)$
$\delta(x\pm x_0,y\pm y_0)$	$\exp[\pm i2\pi(ux_0+vy_0)]$
$\exp[\pm i2\pi(u_0x+v_0y)]$	$\delta(u\mp u_0,v\mp v_0)$
$\exp[\pm i2\pi(u_0x+v_0y)]g(x,y)$	$G(u\mp u_0,v\mp v_0)$
$g(x)h(y)$	$G(u)H(v)$
$g(x)$	$G(u)\delta(v)$
$\delta^{(k,l)}(x,y)$	$(i2\pi u)^k(i2\pi v)^l$
$g^{k,l}(x,y)$	$(i2\pi u)^k(i2\pi v)^lG(u,v)$
x^ky^l	$(-i2\pi)^{-k-l}\delta^{(k,l)}(u,v)$
$\nabla^2 g(x,y)$	$-4\pi^2(u^2+v^2)G(u,v)$
$\int_{-\infty}^{x}\int_{-\infty}^{\infty}g(\xi,\eta)d\xi d\eta$	$\frac{1}{2}\left[G(0,0)\delta(u)+\frac{G(u,0)}{i\pi u}\right]\delta(v)$
$g(x,y)*h(x,y)$	$G(u,v)H(u,v)$
$g(x,y)h(x,y)$	$G(u,v)*H(u,v)$
$g(x,y)\otimes h(x,y)$	$G^{*}(u,v)H(u,v)$
$g(x,y)^{*}h(x,y)$	$G(u,v)\otimes H(u,v)$
$\exp[\pm i\pi(x^2+y^2)]$	$\pm i\exp[\mp i\pi(u^2+v^2)]$
$\exp[-\pi(x^2+y^2)]$	$\exp[-\pi(u^2+v^2)]$
$\mathrm{rect}(x)\mathrm{rect}(y)$	$\mathrm{sinc}(u)\mathrm{sinc}(v)$
$\mathrm{sinc}(x)\mathrm{sinc}(y)$	$\mathrm{rect}(u)\mathrm{rect}(v)$
$\Lambda(x)\Lambda(y)$	$\mathrm{sinc}^2(u)\mathrm{sinc}^2(v)$

1.2 空间带宽积和测不准关系式

1.2.1 空间带宽积与自由度

如果信号 g 在频域内不为零的分量限制在某一区域内，则称为"带限函数". Whittaker-Shannon 抽样定律[1]指出：带限函数 $g(x,y)$ 被它的抽样值的无穷集合 $\{g_{mn}=g(m/\Delta u,n/\Delta v)\}$ 完全确定，式中 $\Delta u,\Delta v$ 是频带的宽度，$m,n=0,\pm1,\pm2,\cdots$. 傅里叶变换及解析函数的一般理论告诉我们：频域内的带限函数，在空域内必然扩展到全平面，因为带限函数的傅里叶变换是一个解析函数，它不可能在一个有限的区域内处处为零，否则通过解析开拓就可以证明这个函数在全平面内处处为零.

然而，用来测量这个信号的实际测量系统的输入平面总是有限制的，设信号被限制在矩形区域 $r\left[-\frac{\Delta x}{2},\frac{\Delta x}{2},-\frac{\Delta y}{2},\frac{\Delta y}{2}\right]$内，又设系统的带宽 $\Delta u,\Delta v$ 与抽样间隙 X,Y 满足倒数的关系，则在 r 内共有抽样点 N 个，

$$N=\frac{\Delta x\Delta y}{XY}=\Delta x\Delta y\Delta u\Delta v=SW, \tag{1}$$

式中

$$S=\Delta x\Delta y, \tag{2}$$

$$W=\Delta u\Delta v, \tag{3}$$

SW 称空间带宽积，是评价系统性能的重要参数，(1)式指出通过系统的样点数等于空间带宽积.

上面的讨论已指出：一个在频域中非无限扩展的信号(带限信号)，在空域中必然是无限扩展的，若用一个具有有限大小的输入端面的系统对该信号进行测量，必然造成信息量的损失，使测量结果失真. 一般来讲信号在空域中不显著为零的分量只分布在有限区域内，例如矩形 r 内，那么这个信号就被它的 N 个样值基本上确定了. 我们称这个信号有 N 个自由度，显然自由度数等于空间

带宽积．如果系统的输入端面的尺寸小于 r，则自由度数将小于 N．所以空间带宽积与其说是信号的特征，还不如说是系统的特征，因为系统有限的空域和频域尺寸限制了通过它的信息量．例如对于一个成像系统，限制空域尺寸的是视场光阑的大小，限制频域尺寸的是孔径光阑的大小．显然视场越大、孔径越大的系统能传递更多的信息．

1.2.2 系统的分辨率

让我们来考虑一个具有低通滤波性能的系统的分辨率，即输入平面上能被系统分辨开来的两个点的最小间距（最小分辨长度）的倒数．

由抽样定理可知，对任意输入信号 $g(x,y)$ 来讲，由于系统频率响应特性的限制，其效果都是带限的（不论信号是否具有带限特性），因此可以用抽样函数 $g_s(x,y)$ 来代替它，尽管 $g_s(x,y)$ 是离散的点阵，但只要抽样点充分稠密，即条件

$$X \leqslant 1/\Delta u, \qquad Y \leqslant 1/\Delta v \tag{4}$$

满足时，对于系统输出端而言，g_s 和 g 等价，在输出端并不能觉察出 g_s 的周期结构，或者说 g_s 包含的脉冲是不可分辨的．

反之，当条件(4)不满足时，g_s 和 g 对于输出端不再等价，从而在输出端就能觉察出 g_s 的周期结构，或者讲 g_s 中两个相邻脉冲能够被系统分辨开来．这样一来，系统的最小分辨长度 δx 和 δy 应当与(4)式表示的 X、Y 同数量级，从而与带宽成反比：

$$\delta x \propto 1/\Delta u, \qquad \delta y \propto 1/\Delta v. \tag{5}$$

最小分辨长度与空间带宽积的关系为

$$\delta x \delta y \propto \Delta x \Delta y / SW, \tag{6}$$

可见在给定输入端面尺寸 Δx，Δy 后，SW 越大，最小分辨长度就越小，系统的分辨率就越高，测量过程的失真越小．(5)和(6)式都可用作系统分辨率的估算．不过这些标准毕竟过于粗糙，例如它们不能回答两个带宽相同的系统的质量有什么差别的问题．为

此，我们进一步讨论等效带宽.

1.2.3 等效带宽和测不准关系

仅考虑一维情况. 函数 $g(x)$ 和它的频谱 $G(u)$ 的关系为

$$G(u) = \int_{-\infty}^{\infty} g(x)\exp(-\mathrm{i}2\pi ux)\mathrm{d}x, \tag{7}$$

$$g(x) = \int_{-\infty}^{\infty} G(u)\exp(\mathrm{i}2\pi ux)\mathrm{d}u, \tag{8}$$

由以上两式可得

$$G(0) = \int_{-\infty}^{\infty} g(x)\mathrm{d}x, \tag{9}$$

$$g(0) = \int_{-\infty}^{\infty} G(u)\mathrm{d}u. \tag{10}$$

进一步设信号在空域和频域中不显著为 0 的分量都集中在原点近旁有限区域内，则我们可以用下面的方法近似度量 $g(x)$ 和 $G(u)$ 的弥散或展宽的程度. 引入 $\bar{s}$ 和 $\widetilde{w}$：

$$\bar{s} = \int_{-\infty}^{\infty} g(x)\mathrm{d}x \Big/ g(0), \tag{11}$$

$$\widetilde{w} = \int_{-\infty}^{\infty} G(u)\mathrm{d}u \Big/ G(0). \tag{12}$$

$\widetilde{w}$ 有明显的意义：如一个矩形高度等于 $G(0)$，面积与曲线 $G(u)$ 下的面积相同，则它的宽度为 $\widetilde{w}$，$\widetilde{w}$ 又称为“等效带宽”，Goodman[1] 提出了等效带宽的概念，它是频谱曲线展宽程度的某种度量，参见图 1.1，$G(u)$ 越宽，$\widetilde{w}$ 越大，因而常用来评价系统的性能. 宋菲君[5] 讨论了小像差光学系统的等效带宽和测不准问题的关系.

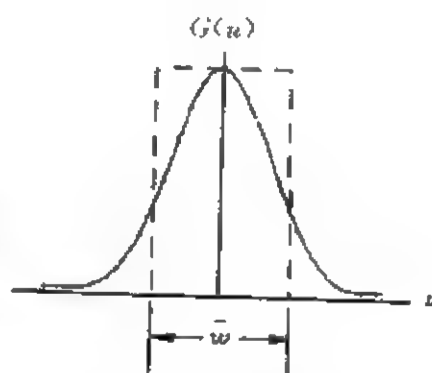


图 1.1 等效带宽
矩形下的面积与 $G(u)$ 下的面积相等，矩形的宽度为等效带宽

将(11)、(12)交叉相除得到

$$\tilde{s}\tilde{w} = 1. \tag{13}$$

由于 $\tilde{s}$ 可以表征信号在空域的展宽或弥散，上式意味着信号在空域和频域中的展宽是互相制约的.

假设我们要对信号进行长度或位置测量，测量系统可看成是对被测对象的一个变换，在位置测量时必须使系统首先“对准”空间的一个定点或长度的一个端点，该点可以用 δ 函数表示，它就是系统的输入，而输出恰恰就是系统的脉冲响应 h. 必须指出，通过测量我们只能获得 h 所包含的信息，我们永远无法直接得到被测点本身. 所有测量系统的等效带宽 $\tilde{w}$ 都是有限的，从而 δ 函数的脉冲响应 h 就有一定的弥散 $\tilde{s}$，它表征了对准误差，因而也就是系统空间分辨率大小的度量. 注意到 $\tilde{w}$ 取决于整个频谱函数 $G(u)$，因此两个系统即使有等同的截止频率，由于 $G(u)$ 不相同，也会得到不同的等效带宽 $\tilde{w}$，因而 $\tilde{s}$ 也不一致. 一般来讲，$\tilde{w}$ 越大，频响特性就越好，脉冲响应的弥散 $\tilde{s}$ 就越小. 由于 $\tilde{w}=\infty$ 的系统不存在，所以 $\tilde{s}$ 永远不等于 0. 在这个意义上讲，测量永远都不是绝对准确的，(13)式称为光学系统的测不准关系，它与量子力学中的测不准关系实质上一致.

1.2.4 广义测不准关系的严格推导[3]

首先我们来定义一个信号 $g(x)$ 空域分布的中心

$$x_c = \frac{(g(x), xg(x))}{(g(x), g(x))} = \frac{\int_{-\infty}^{\infty} g^*(x)xg(x)\mathrm{d}x}{\int_{-\infty}^{\infty} g^*(x)g(x)\mathrm{d}x}, \tag{14}$$

式中(·，·)表示函数 g 和 h 的内积. 其次，信号 $g(x)$ 的空域宽度 Δx 定义为

$$(\Delta x)^2 = \frac{(g(x),(x-x_c)^2 g(x))}{(g(x),g(x))} = \frac{\int_{-\infty}^{\infty} g^*(x)(x-x_c)^2 g(x)\mathrm{d}x}{\int_{-\infty}^{\infty} g^*(x)g(x)\mathrm{d}x}. \tag{15}$$

信号 g 的傅里叶频谱 $G(u)$ 在频域中的中心 u_c 和宽度 Δu 可仿此定义：

$$u_c = \frac{(G(u),uG(u))}{(G(u),G(u))} = \frac{\int_{-\infty}^{\infty} G^*(u)uG(u)\mathrm{d}u}{\int_{-\infty}^{\infty} G^*(u)G(u)\mathrm{d}u}, \tag{16}$$

$$(\Delta u)^2 = \frac{(G(u),(u-u_c)^2 G(u))}{(G(u),G(u))} = \frac{\int_{-\infty}^{\infty} G^*(u)(u-u_c)^2 G(u)\mathrm{d}u}{\int_{-\infty}^{\infty} G^*(u)G(u)\mathrm{d}u}. \tag{17}$$

上述定义并不是对所有的信号都有意义，它要求信号在空域和频域的分布都相对集中，或者说当宗量的绝对值趋于∞时 $g(x)$ 和 $G(u)$ 的衰减足够快.

下面我们来证明信号空域宽度 Δx 和频域宽度 Δu 之积有一个下限，即

$$\Delta x\Delta u \geqslant 1/4\pi, \tag{18}$$

首先我们来证明下面的关系式

$$\int_{-\infty}^{\infty} u^2 G^*(u)G(u)\mathrm{d}u = \frac{1}{4\pi^2}\int_{-\infty}^{\infty} g'(x)g'^*(x)\mathrm{d}x, \tag{19}$$

上式左边积分内的 $G^*(u)$ 和 $G(u)$ 可以用傅里叶变换表示：

$$\begin{aligned}\int_{-\infty}^{\infty} u^2 G^*(u)G(u)\mathrm{d}u &= \int_{-\infty}^{\infty}\mathrm{d}u\left[u\int_{-\infty}^{\infty} g^*(x)\exp(\mathrm{i}2\pi ux)\mathrm{d}x\right] \\ &\quad\times\left[u\int_{-\infty}^{\infty} g(\xi)\exp(-\mathrm{i}2\pi u\xi)\mathrm{d}\xi\right] \\ &= \int_{-\infty}^{\infty} I_1(u)I_2(u)\mathrm{d}u,\end{aligned} \tag{20}$$

式中 I_1 可以用分部积分方法来改写：

$$I_1(u)=u\int_{-\infty}^{\infty}g(\xi)\exp(-\mathrm{i}2\pi u\xi)\mathrm{d}\xi$$

$$=-\frac{1}{\mathrm{i}2\pi}\int_{-\infty}^{\infty}g(\xi)\mathrm{d}[\exp(-\mathrm{i}2\pi u\xi)]$$

$$=-\frac{1}{\mathrm{i}2\pi}\Big[g(\xi)\exp(-\mathrm{i}2\pi u\xi)\Big|_{-\infty}^{\infty}$$

$$-\int_{-\infty}^{\infty}g'(\xi)\exp(-\mathrm{i}2\pi u\xi)\mathrm{d}\xi\Big]$$

$$=\frac{1}{\mathrm{i}2\pi}\int_{-\infty}^{\infty}g'(\xi)\exp(-\mathrm{i}2\pi u\xi)\mathrm{d}\xi, \qquad (21)$$

在推导中已假设 $g(\xi)\exp(-\mathrm{i}2\pi u\xi)$ 在 $\xi\to\infty$ 时为零. 类似地，我们有

$$I_2(u)=-\frac{1}{\mathrm{i}2\pi}\int_{-\infty}^{\infty}g^*(x)\mathrm{d}[\exp(\mathrm{i}2\pi ux)]$$

$$=-\frac{1}{\mathrm{i}2\pi}\int_{-\infty}^{\infty}g'^*(x)\exp(\mathrm{i}2\pi ux)\mathrm{d}x. \qquad (22)$$

以(21)，(22)代入(20)，交换积分次序，并运用 δ 函数的积分表达式就得到(19)式.

我们再引入下面形式的 Schwarz 不等式：

$$4\int gg^*\mathrm{d}x\int hh^*\mathrm{d}x\geqslant\left|\int(g^*h+gh^*)\mathrm{d}x\right|^2, \qquad (23)$$

运用(19)，(23) 式及节 1.1(26)式的结果，我们有

$$(\Delta x)^2(\Delta u)^2=\frac{\int_{-\infty}^{\infty}x^2gg^*\mathrm{d}x\int_{-\infty}^{\infty}u^2GG^*\mathrm{d}u}{\int_{-\infty}^{\infty}gg^*\mathrm{d}x\int_{-\infty}^{\infty}GG^*\mathrm{d}u}$$

$$=\frac{\int_{-\infty}^{\infty}xg\cdot xg^*\mathrm{d}x\int_{-\infty}^{\infty}g'g'^*\mathrm{d}x}{4\pi^2\left(\int_{-\infty}^{\infty}gg^*\mathrm{d}x\right)^2}$$

$$\geqslant \frac{\left|\int_{-\infty}^{\infty}(xg^* \cdot g' + xg \cdot g'^*)\mathrm{d}x\right|^2}{4\pi^2\left(\int_{-\infty}^{\infty} gg^*\mathrm{d}x\right)^2}, \tag{24}$$

其中分子中的积分可化为

$$\int_{-\infty}^{\infty} x\frac{\mathrm{d}}{\mathrm{d}x}(gg^*)\mathrm{d}x = xgg^*\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} gg^*\mathrm{d}x = -\int_{-\infty}^{\infty} gg^*\mathrm{d}x, \tag{25}$$

在推导中假定 xgg^* 当 $|x|\to\infty$ 时为零. 以(25)式代入(24)式，得到

$$(\Delta x)^2(\Delta u)^2 \geqslant 1/16\pi^2$$

即(18)式，可称为广义的测不准关系式.

二维的情况可仿此推导.

$\Delta x\Delta u$ 的值取决于信号函数的具体形式. 当 $g(x)$ 取高斯函数时，即

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right), \tag{26}$$

$$\Delta x = \sigma/\sqrt{2}, \tag{27}$$

与(26)式对应的频域信号为

$$G(u) = \exp(-2\pi^2\sigma^2u^2), \tag{28}$$

它的宽度可由(17)式算出：

$$\Delta u = 1/2\sqrt{2}\pi\sigma, \tag{29}$$

从而有

$$\Delta x\Delta u = 1/4\pi, \tag{30}$$

亦即对于高斯信号而言，(18)式中的等号成立，在第六章“小波变换”中，我们将看到这一例子的意义.

1.3 平面波的角谱和角谱的衍射

本节讨论衍射问题. 我们当然可以从基尔霍夫条件和瑞利-

索末菲衍射公式出发讨论这一问题．但从方法论上来看，这样的做法依然没有脱离传统的物理光学的框架．下面我们完全沿着另一途径，即从变换光学入手来讨论衍射效应．

1.3.1 角谱

设单色光波沿 z 方向传播，照射到 xy 平面上，在 xy 平面上的光场复振幅分布用函数 $\psi(x,y,0)=\psi(x,y)$ 表示．我们的目的是要寻求与 xy 平面相距为 z 且与 xy 平面平行的下一个平面上的光场分布．$\psi(x,y)$ 可进行频谱分析如下：

$$\psi(x,y,0) = \psi(x,y) = \iint_{-\infty}^{\infty} A(u,v)\exp[-\mathrm{i}2\pi(ux+vy)]\mathrm{d}u\mathrm{d}v. \tag{1}$$

一个波矢量为 $\boldsymbol{k}$ 的平面波

$$\varphi_0(x,y,z) = \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}) = \exp\left[\mathrm{i}\,\frac{2\pi}{\lambda}(\alpha x+\beta y+\gamma z)\right], \tag{2}$$

其中 α、β 和 γ 是 $\boldsymbol{k}$ 的方向余弦．引入二维矢量

$$\boldsymbol{\alpha} = (\alpha,\beta), \tag{3}$$

则在 $z=0$ 的平面上

$$\varphi_0(x,y,0) = \exp\left[\mathrm{i}\,\frac{2\pi}{\lambda}(\boldsymbol{\alpha}\cdot\boldsymbol{r})\right] = \exp\left[\mathrm{i}\,\frac{2\pi}{\lambda}(\alpha x+\beta y)\right]. \tag{4}$$

将(4)式和(1)式作比较，发现只要取

$$u = \alpha/\lambda, \qquad v = \beta/\lambda, \tag{5}$$

则(1)式可用 $\boldsymbol{\alpha}$ 表示为

$$\psi(x,y) = \iint_{-\infty}^{\infty} A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\right)\exp\left[\mathrm{i}2\pi\left(\frac{\alpha}{\lambda}x+\frac{\beta}{\lambda}y\right)\right]\mathrm{d}\left(\frac{\alpha}{\lambda}\right)\mathrm{d}\left(\frac{\beta}{\lambda}\right), \tag{6}$$

(6)式表示：$z=0$ 平面上的场，或者说，透过 xy 平面向 $+z$ 方向传播的波，可以用不同方向的平面波展开．(5)式表示复振幅分布的空间频率正比于 α/λ 或 β/λ，在 $\psi(x,y)$ 中的低频分量对应于与 z

轴夹角不大的平面波分量．而高频分量则对应于与 z 轴夹角较大的平面波分量，这是一个重要的概念．不同方向的平面波的权函数 $A(\alpha/\lambda,\beta/\lambda)$ 称为 $\psi(x,y)$ 的角谱，它和空间频谱的实质是相同的．

$A(\alpha/\lambda,\beta/\lambda)$ 与 $\psi(x,y)$ 的关系就是傅里叶变换：

$$A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\right)=\iint_{-\infty}^{\infty}\psi(x,y)\exp\left[-\mathrm{i}2\pi\left(\frac{\alpha}{\lambda}x+\frac{\beta}{\lambda}y\right)\right]\mathrm{d}x\mathrm{d}y. \tag{7}$$

(6)和(7)两式构成傅里叶变换对．

1.3.2 角谱的传播

现在来进一步寻求 $\psi(x,y,z)$ 与 $\psi(x,y,0)$ 的关系．为此，首先求 $A(\alpha/\lambda,\beta/\lambda;z)$ 与 $A(\alpha/\lambda,\beta/\lambda)$ 的关系．$A(\alpha/\lambda,\beta/\lambda;z)$ 与 $\psi(x,y,z)$ 的关系显然为

$$A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z\right)=\iint_{-\infty}^{\infty}\psi(x,y,z)\exp\left[-\mathrm{i}2\pi\left(\frac{\alpha}{\lambda}x+\frac{\beta}{\lambda}y\right)\right]\mathrm{d}x\mathrm{d}y, \tag{8}$$

$$\psi(x,y,z)=\iint_{-\infty}^{\infty}A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z\right)\exp\left[\mathrm{i}2\pi\left(\frac{\alpha}{\lambda}x+\frac{\beta}{\lambda}y\right)\right]\mathrm{d}\left(\frac{\alpha}{\lambda}\right)\mathrm{d}\left(\frac{\beta}{\lambda}\right), \tag{9}$$

以 $\psi(x,y,z)$ 代入亥姆霍兹方程，交换积分与微分的次序，可知 $A(\alpha/\lambda,\beta/\lambda;z)$ 也满足亥姆霍兹方程：

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}z^2}+k_z^2\right)A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z\right)=0, \tag{10}$$

式中

$$k_z=\frac{2\pi}{\lambda}\sqrt{1-(\alpha^2+\beta^2)}. \tag{11}$$

(10)式的一个解是

$$A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z\right)=A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\right)\exp(\mathrm{i}k_z z)$$

$$= A\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right)\exp\left[\mathrm{i}\,\frac{2\pi z}{\lambda}\,\sqrt{1-(\alpha^2+\beta^2)}\right]. \tag{12}$$

当 $\alpha^2+\beta^2<1$ 时，光波沿 $+z$ 方向传播的效果，在频域内表现为乘以一个沿 z 轴的相位延迟因子 $\exp\left[\mathrm{i}\,\frac{2\pi z}{\lambda}\sqrt{1-(\alpha^2+\beta^2)}\right]$. 在第二章，我们将看到这一效应等价于空间滤波.

当 $\alpha^2+\beta^2\geqslant 1$ 时，取正数 $\mu=\sqrt{\alpha^2+\beta^2-1}$，则角谱

$$A(\alpha/\lambda,\beta/\lambda;z) = A(\alpha/\lambda,\beta/\lambda)\exp(-2\pi\mu z/\lambda), \tag{13}$$

表示一个随 z 的增大迅速衰减的波，称隐失波，它只存在于很接近于 xy 平面的一个薄层内，这是近场光学要讨论的问题. 下面我们只考虑前一种情况.

1.3.3 菲涅耳衍射

将(12)式中相因子内的根号作泰勒展开：

$$\sqrt{1-\lambda^2\rho^2} = \frac{1}{2}\lambda^2\rho^2 - \frac{1}{8}\lambda^4\rho^4 + \cdots \tag{14}$$

式中 $\rho^2=u^2+v^2$. 在上式中只保留二级小量，则

$$\begin{aligned} A(\alpha/\lambda,\beta/\lambda;z) &= A(u,v;z) \\ &= A(u,v)\exp(\mathrm{i}2\pi z/\lambda) \\ &\quad \times \exp[-\mathrm{i}\pi\lambda z(u^2+v^2)], \end{aligned} \tag{15}$$

由于

$$A(u,v)\Leftrightarrow\psi(x,y),$$

$$\exp[-\mathrm{i}\pi\lambda z(u^2+v^2)]\Leftrightarrow\exp[\mathrm{i}\pi(x^2+y^2)/\lambda z]/(\mathrm{i}\lambda z),$$

根据卷积的变换性质，相应的空域信号为

$$\begin{aligned} \psi(x,y,z) &= \frac{1}{\mathrm{i}\lambda z}\exp(\mathrm{i}2\pi z/\lambda)\psi(x,y) * \exp[\mathrm{i}\pi(x^2+y^2)/\lambda z] \\ &= \frac{1}{\mathrm{i}\lambda z}\exp(\mathrm{i}2\pi z/\lambda)\iint_{-\infty}^{\infty}\psi(\xi,\eta) \\ &\quad \times \exp\{\mathrm{i}\pi[(x-\xi)^2+(y-\eta)^2]/\lambda z\}\mathrm{d}\xi\mathrm{d}\eta, \end{aligned} \tag{16}$$

式中 $\psi(\xi,\eta)$ 表示 $z=0$ 的光场复振幅分布. (16)式是菲涅耳衍射

的公式. 上述积分在 $z=0$ 的平面进行.

1.3.4 夫琅和费衍射

在(16)式中加入更为强烈的近似条件

$$z \gg \pi(\xi^2 + \eta^2)/\lambda, \tag{17}$$

则该式化为

$$\psi(x,y,z) = \frac{1}{i\lambda z}\exp(i2\pi z/\lambda)\exp[i\pi(x^2 + y^2)/\lambda z]$$

$$\times \iint_{-\infty}^{\infty}\psi(\xi,\eta)\exp[-i2\pi(\xi x + \eta y)/\lambda z]d\xi d\eta, \tag{18}$$

就化为远场衍射即夫琅和费衍射的情况. (18)式还可表为

$$\psi(x,y,z) = \frac{A}{\lambda z}\Psi\left(\frac{x}{\lambda z},\frac{y}{\lambda z}\right). \tag{19}$$

上式表示除了与积分变量无关的相位因子 A 以外,Ψ 为 ψ 的傅里叶变换,频域宗量为 $x/\lambda z$ 及 $y/\lambda z$.

1.3.5 角谱的衍射

设在 xy 平面上有一个不透光的屏,屏上带一个透光的孔,孔的复数透过率用光瞳函数 $p(x,y)$ 来表示,$p(x,y)$ 可以是复数. 这样一来,屏后面的透射场 ψ_t 可用入射波的场 ψ_i 表为

$$\psi_t(x,y) = \psi_i(x,y)p(x,y), \tag{20}$$

在频域中,上式变为

$$A_t(\alpha/\lambda,\beta/\lambda) = A_i(\alpha/\lambda,\beta/\lambda) * P(\alpha/\lambda,\beta/\lambda), \tag{21}$$

式中 P 为 p 的角谱. (21)式说明透射波的角谱为入射波的角谱与光瞳函数角谱的卷积. 引入光阑后,一般来讲信号的空间分布受到压缩. 根据测不准原理,信号在频域中的分布必然展宽. (21)式所示的卷积运算的结果,总是使入射波的角谱变得更加平滑,换言之,有更多的能量扩散到高频段中去.

(12)式为角谱在自由空间中的衍射公式. 如果考虑到 xy 平

面上光瞳函数的作用，(12)式改写为

$$A(\alpha/\lambda,\beta/\lambda;z)=[A_i(\alpha/\lambda,\beta/\lambda;z)*P(\alpha/\lambda,\beta/\lambda)]$$
$$\times\exp[i2\pi z\sqrt{1-(\alpha^2+\beta^2)}/\lambda]. \quad (22)$$

(12)式或(22)式原则上可以解决任何光波的传播及衍射问题.

1.4 透镜系统的傅里叶变换性质

1.3 节公式(18)表明，远场衍射具有傅里叶变换的特性. 由于薄透镜或透镜组的后焦面等价于∞，因而可以想像凡是具有正焦距的光学系统都应当具有傅里叶变换的功能.

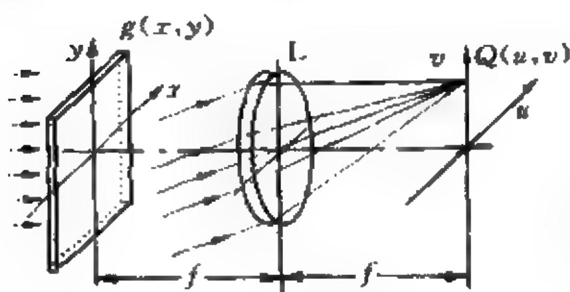


图 1.2 透镜组聚焦

如图 1.2 所示，设用振幅为 1 的单色平面波照射一个在 xy 平面上，且振幅透过率为 $g(x,y)$ 的物体，则物体后面的场为 $g(x,y)$. 将它置于透镜组 L 的前焦面上，根据上一节中一个光场展开成为平面波角谱的做法，我们把 g 展开：

$$g(x,y)=\iint_{-\infty}^{\infty}G\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\right)\exp\left[i2\pi\left(\frac{\alpha}{\lambda}x+\frac{\beta}{\lambda}y\right)\right]d\frac{\alpha}{\lambda}d\frac{\beta}{\lambda}. \quad (1)$$

由于透镜组具有聚焦的特性，所有方向相同，即具有同样的方向余弦 α,β 的入射波都将会聚到透镜组后焦面的一点 $Q(u,v)$ 上. 当透镜组焦距 $f\gg\sqrt{u^2+v^2}$ 时，即 Q 点很接近于原点时，有下面的近似等式

$$u=\alpha f,\qquad v=\beta f, \quad (2)$$

$g(x,y)$ 的角谱中所有方向余弦为 α,β 的角谱分量都对 Q 点有贡献，Q 点的的复振幅自然就等于 $G(\alpha/\lambda,\beta/\lambda)$，因而后焦面上的复振幅分布为

$$G(\alpha/\lambda,\beta/\lambda)=G(u/\lambda f,v/\lambda f). \quad (3)$$

这样一来，透镜组的后焦面就成为信号的频域，透镜组起了傅里叶变换的作用. 进一步，大部分具有聚焦性能的器件，例如反光镜、自聚焦透镜等，都具有傅里叶变换的功能，薄透镜的傅里叶变换功能可以直接计算出来，但它只是光学傅里叶变换器件的一个特例. 在本书中，我们用 u，v 来表示频域的坐标，在不会引起混淆之处，u，v 也用以表示空间频率变量. 在一维的情形下也用 ν 来表示空间频率变量.

注意(2)式只是近轴近似的结果. 严格来说，

$$u = f\tan\theta = f\alpha / \sqrt{1-\alpha^2}, \tag{4}$$

式中 θ 是波矢量 $\boldsymbol{k}$ 与 z 轴的夹角. 为简单起见，设 $\boldsymbol{k}$ 位于 xz 平面内，参见图 1.3. (4)式又称正切条件，只是在 θ 很小时，才满足(2)式. 当 θ 较大时，傅里叶平面(后焦面)上的线度 u 与空间频率 α/λ 并不满足正比关系.

此外，从几何光学我们还知道，一个像差校正得很好的透镜必须满足正弦条件[4]，而正弦条件与正切条件是难以同时满足的，从这个意义上说，性能完善的傅里叶变换透镜是很难设计的.

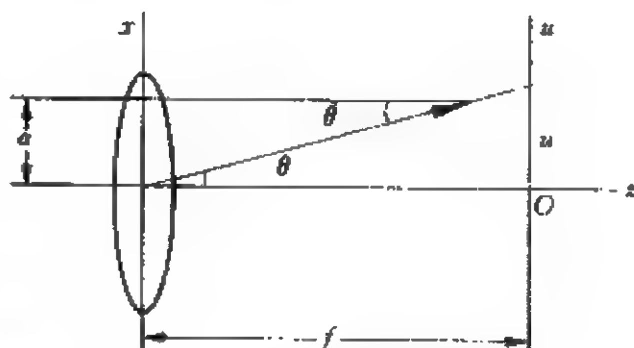


图 1.3 正切条件：在 θ 较小时，$u \approx f \cdot \alpha$

不过在大多数情况下，光学变换是作为近似的模拟变换而加以应用的，再说推导薄透镜的相位变换公式时已经引入了近轴近似. 在大多数应用中，无论是薄透镜或是透镜组仍然是最方便、廉价的光学傅里叶变换器件.

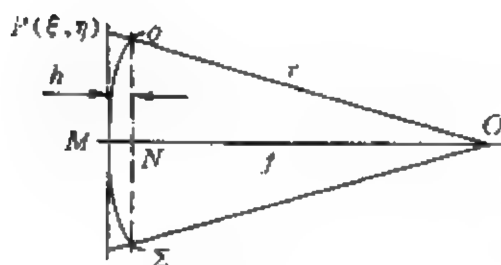


图 1.4

由于透镜系统能够将一个

平面波转换成一个球面波，所以它的相位变换效应可以表为

$$t_l = e^{ik\Delta}e^{-ikr}/r, \tag{5}$$

式中 Δ 为透镜组的等效厚度．e^{-ikr}/r 正是会聚球面波的表达式，$r=OP$（参见图 1.4），O 是会聚球面波的中心，也是透镜系统的焦点，$OQ=OM=f$，f 为焦距，在近轴近似下，$PQ\approx MN=h$，所以我们有

$$r = f + PQ \approx f + h \approx f - (\xi^2 + \eta^2)/2f, \tag{6}$$

式中(ξ,η)是 P 点坐标，代入(5)式，取分母上的 $r\approx f$，我们得到透镜系统的相位变换公式

$$t_l = \frac{1}{f}\exp[(ik(\Delta - f)]\exp\left(-ik\frac{\xi^2+\eta^2}{2f}\right), \tag{7}$$

下面我们来推导透镜系统对图像的变换公式．

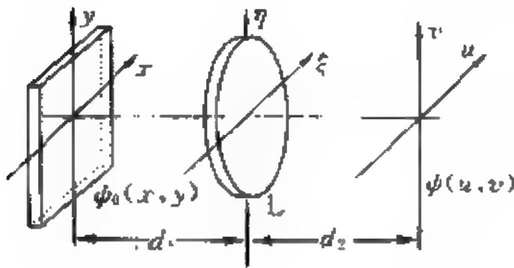


图 1.5　透镜的变换效应

参见图 1.5，设输入平面的透过率为 $\psi_0(x,y)$，它位于透镜 L 前 d_1 处．输出平面 uv 位于 L 后 d_2 处．物体用振幅为 1 的单色光波照明．

设光波在 d_1 和 d_2 范围内的传播满足菲涅耳近似条件，则由 1.3 节公式(16)透镜前表面的场 ψ_l 可表为

$$\psi_l(\xi,\eta)=\frac{e^{ikd_1}}{i\lambda d_1}\iint_{-\infty}^{\infty}\psi_0(x,y)\exp\left\{i\frac{k}{2d_1}[(\xi-x)^2+(\eta-y)^2]\right\}dxdy, \tag{8}$$

透镜 L 的相位变换效应可表为

$$\psi_l'(\xi,\eta) = t_l\psi_l = \frac{1}{f}e^{-ikf}\exp\left[i\frac{k}{2f}(\xi^2+\eta^2)\right]\psi_l(\xi,\eta), \tag{9}$$

在其中我们略去了常数相位项 $\exp(ik\Delta)$．

再用一次菲涅耳变换的公式，得到输出平面(u,v)上的场

$$\psi(u,v)=\frac{e^{ikd_2}}{i\lambda d_2}\iint_{-\infty}^{\infty}\psi_l'(\xi,\eta)\exp\left\{i\frac{k}{2d_2}[(u-\xi)^2+(v-\eta)^2]\right\}d\xi d\eta. \tag{10}$$

将(8),(9)代入(10)式,得

$$\psi(u,v)=-\frac{1}{\lambda^2 d_1 d_2 f}e^{ik(d_1+d_2-f)}\exp\left[i\frac{k}{2d_2}(u^2+v^2)\right]$$
$$\times\iint_{-\infty}^{\infty}\psi_0(x,y)\exp\left[i\frac{k}{2d_1}(x^2+y^2)\right]I(x,y)dxdy, \tag{11}$$

其中

$$\begin{aligned}I(x,y)=&\iint_{-\infty}^{\infty}\exp\left\{i\frac{k}{2}\left[\left(\frac{1}{d_1}+\frac{1}{d_2}-\frac{1}{f}\right)(\xi^2+\eta^2)\right.\right.\\&\left.\left.-2\left(\frac{x}{d_1}+\frac{u}{d_2}\right)\xi-2\left(\frac{y}{d_1}+\frac{v}{d_2}\right)\eta\right]\right\}d\xi d\eta\\&=I_1(x,y)I_2(x,y),\end{aligned} \tag{12}$$

$$I_j=\int_{-\infty}^{\infty}\exp\left[i\left(\frac{k\varepsilon}{2}\xi^2-k\rho_j\xi\right)\right]d\xi\quad(j=1,2), \tag{13}$$

$$\left.\begin{aligned}&\varepsilon=1/d_1+1/d_2-1/f,\\&\rho_1=x/d_1+u/d_2,\\&\rho_2=y/d_1+v/d_2.\end{aligned}\right\} \tag{14}$$

以下分两种情况:

(a) $\varepsilon\neq 0$

$$I_1=\sqrt{\frac{i\lambda}{\varepsilon}}\exp\left[-i\frac{k}{2\varepsilon}\left(\frac{x}{d_1}+\frac{u}{d_2}\right)\right], \tag{15}$$

$$I_2=\sqrt{\frac{i\lambda}{\varepsilon}}\exp\left[-i\frac{k}{2\varepsilon}\left(\frac{y}{d_1}+\frac{v}{d_2}\right)\right], \tag{16}$$

其中用到积分公式

$$\int_{-\infty}^{\infty}\exp[i(A\xi^2+B\xi)]d\xi=\sqrt{\frac{i\pi}{A}}\exp\left(-i\frac{B^2}{4A}\right)\quad(A\neq 0). \tag{17}$$

以(15),(16)式代入(12)式,再代入(11)式,经整理,得到

$$\psi(u,v)=\frac{e^{ik(d_1-d_2-f)}}{i\lambda\varepsilon d_1d_2f}\exp\left[i\frac{k}{2\varepsilon d_1d_2}\left(1-\frac{d_1}{f}\right)(u^2+v^2)\right]$$

$$\times\iint_{-\infty}^{\infty}\psi_0(x,y)\exp\left\{i\frac{k}{2\varepsilon d_1d_2}\left[\left(1-\frac{d_2}{f}\right)(x^2+y^2)\right.\right.$$

$$\left.\left.-2(xu+yv)\right]\right\}dxdy, \tag{18}$$

当 $d_2=f$,即以后焦面作为输出平面,则(18)式化作

$$\psi(u,v)=\frac{e^{ikd_1}}{i\lambda f^2}\exp\left[i\frac{k}{2f}\left(1-\frac{d_1}{f}\right)(u^2+v^2)\right]$$

$$\times\iint_{-\infty}^{\infty}\psi_0(u,v)\exp\left[-i\frac{2\pi}{\lambda f}(xu+yv)\right]dxdy, \tag{19}$$

此时 ψ 是 ψ_0 的傅里叶变换(相位因子除外,在探测光强时相位因子不起作用),宗量是$(u/\lambda f,v/\lambda f)$.

当 $d_1=f$ 时,相位因子消去,

$$\psi(u,v)=\frac{e^{ikf}}{i\lambda f^2}\iint_{-\infty}^{\infty}\psi_0(u,v)\exp\left[-i\frac{2\pi}{\lambda f}(xu+yv)\right]dxdy, \tag{20}$$

在一般情况下,d_1 和 d_2 与 f 并不相等. 在这种情况下,有可能实现广义傅里叶变换(分数阶傅里叶变换). 我们将在第五章中详细讨论这一课题.

(b) $\varepsilon=0$,即输入输出平面关于透镜组满足成像关系. 此时

$$I_1=\lambda d_1\delta(x-u/\beta), \tag{21}$$

$$I_2=\lambda d_2\delta(y-v/\beta), \tag{22}$$

式中 $\beta=-d_2/d_1$ 为系统的横向放大率. 以(21)、(22)式代入(12)、(11)式得

$$\psi(u,v)=\frac{e^{ik(d_1+d_2-f)}}{\beta f}\exp\left[-\frac{i\pi}{\lambda\beta f}(u^2+v^2)\right]\psi_2\left(\frac{u}{\beta},\frac{v}{\beta}\right). \tag{23}$$

在输出平面上得到放大 β 倍的像,回到几何光学的结果. 这个结论只是近似成立的,因为我们完全不考虑光瞳函数的影响,也忽略

了透镜的像差.

在本章中,我们用较少篇幅,介绍了二维傅里叶变换和傅里叶光学的主要内容及结论,它们是经典和近代光学信息处理的基础.对以上内容有兴趣的读者,可参考Goodman[1]和Bracewell[3]的经典著作.

参考文献

[1] J. W. Goodman, Introduction to Fourier optics, McGRAW-HILL, New York, 1968.

[2] 羊国光, 宋菲君, 高等物理光学, 中国科技大学出版社, 1991.

[3] R. N. Bracewell, The Fourier Transform and Its Applications, McGRAW-HILL, New York, 1978.

[4] 王之江, 光学设计理论基础, 科学出版社, 1962.

[5] 宋菲君, 光学系统的测不准关系和等效带宽, 光学学报, 6(1986), 137.

第二章 经典光学信息处理

2.1 引 言

能量和质量是基本的物理量,时间和空间是运动物质存在的基本形式,表现能量时-空分布的图形图像则产生了信息. 根据这一定义,任何物体都持续地产生新的信息. 能量从能量源传递到探测器,在能量传递过程中伴随着信息的传递,就形成信号. 探测到的能量中所包含的不需要的信息则称为噪声. 人类是信息的最终探测者,并将信息翻译成抽象的知识.

Yu[1]认为,光不仅是支持生命的能量流的一部分,并且是信息的重要源泉. 一般来说,光学信息可以认为是光波所荷载的信息. 如上所述,信息是以能量起伏分布形成的图像的形式而表现出来的,本书中,光学信息特指光强分布所形成的图像,它可以是日常生活中司空见惯的自然图像,也可以是人造的或人工模拟的图像. 光学信息处理,指的是光学图像的产生、传递、探测和处理,所需要的图像称为信号,在处理过程中伴生的不需要的图像称噪声. 本章介绍经典的光学信息处理,被处理的图形是真实物体的像.

2.2 早 期 发 展

光学信息处理的历史可以一直追溯到125年前阿贝所做的研究工作[3]. 当时他对显微术的理论很有兴趣,打算建立设计显微镜的科学方法. 他发现显微物镜的孔径越大,分辨率越高. 他还发现显微镜所观察的标本中与光波波长同量级的小物体引起光波

的衍射，使光线的传播偏离几何光学的路径．如果物镜的孔径不足以收集物体所有的衍射光，显微镜中物体的像将不是原来物体的完善的像，而是由物镜收集的那一部分衍射光生成的“虚拟像”．根据第一章关于角谱的讨论，可知丢失的信息是高频信息．因此显微镜可以看作是一个低通滤波器．显微镜的孔径越大（严格地讲，应当说是“相对孔径”越大），系统的通频带越宽，物体中所包含的高频信息在成像过程中的损失就越少．光学系统或光学仪器是传递图像信息的系统，在传递过程中信息损失越少，像与物的差别就越小，像的质量就越高．

反之，孔径（相对孔径）越小，在传递过程中高频信息的损失就越大、像的失真或畸变就越严重，清晰度或分辨率越低．

1960 年，Porter[4]用一系列实验阐明了阿贝的理论．一组很细的网状物体 O 用准直的相干光照明，在成像透镜 L 的焦面上出现了网状物体的衍射图形，而在像平面 I 上则出现网状物体的像．如果在焦平面上加各种滤波器，输出图像就会发生变化．例如，在焦平面上放置一个水平的狭缝 S，网格就会变成一组垂直的线条，如图 2.1 所示．

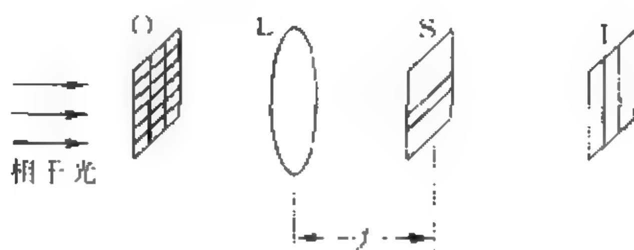


图 2.1　在焦平面上的一个水平狭缝消除了网格像的水平线条

Zernike[5]在 1935 年发明的相衬显微镜对光学信息处理作出了举足轻重的贡献．众所周知，人们看见一个物体，是因为该物体与周围环境有显著的差别，这一类结构称为“振幅”物体，因为该物体调制了照明光的振幅．然而另一类“相位”物体调制了照明光束的相位．由于人眼只能探测光强，不能辨别不同的相位，相位物

体是看不见的. 许多显微镜的标本属于相位物体,因而有必要识别相位物体. 相位物体也能使光波发生衍射,在透镜后焦面上形成衍射像. 如不采取特别的措施,相位物体的像依然是相位物体或相位分布,仍旧看不见. 然而,如果用可变的相位片改变衍射图像,最终的像必然发生改变,这个像不再是相位变化,而包含了可见的振幅变化. 如图 2.2 所示,采用适当厚度的相位片,可以使最终的像的振幅变化与相位变化成正比. 图中 P 为相位物体,L 为透镜,PF 为相位片,AI 为振幅图像,P 用相干光照明.

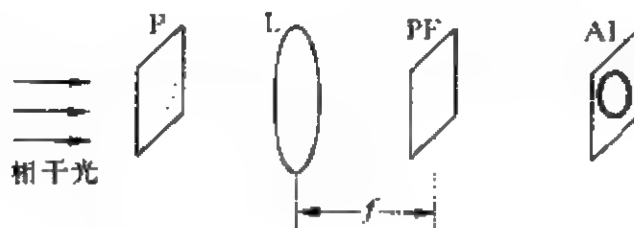


图 2.2 相衬法使相位物体变为可见

事实上,早在 1864 年在阿贝提出他的理论以前,Toepler[6]就发明了相衬法,这一技术称为 Schlieren 方法,早先用来探测透镜的疵病. Schlieren 在德语中是条纹的意思,它不是人的名字. 与上述相衬显微镜类似,在这一方法中,只是简单地把衍射图形挡去一半多一点,透镜中的疵病等相位物体就可以看见. 事实上,这种简单而有效的方法一直沿用至今,使风洞中的气压分布变成可见的图像. 在图 2.3 中 HS 是光阑,它挡去一半多一点的衍射图形. P 仍用相干光照明.

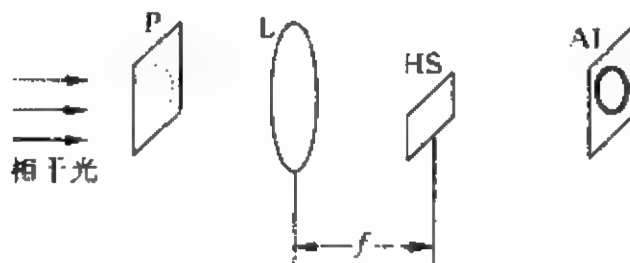


图 2.3 Schlieren 方法同样使相位物体变为可见

2.3 傅里叶处理器

在本世纪 40 年代晚期和 50 年代，发现了把相干光技术运用于非显微图像处理的可能性. 这些早期的研究工作是由 Duffieux[7]，Marechal[8]，Hopkins[9]和 O'Neil[10]进行的. 1960 年，Cutrona 等[11]明确提出用透镜进行傅里叶变换的方案，从而傅里叶变换就可以用透镜简单地实现了. 这一技术的特点，除了构成傅里叶谱以外，还在于系统可以串联运用.

节 1.4 详细论述了透镜傅里叶变换效应. 如图 2.4 所示，透镜 L 后焦面(uv 平面)上的复振幅分布正是位于它前焦面(xy 平面)物体的傅里叶变换. 注意所使用的光必须是相干光.

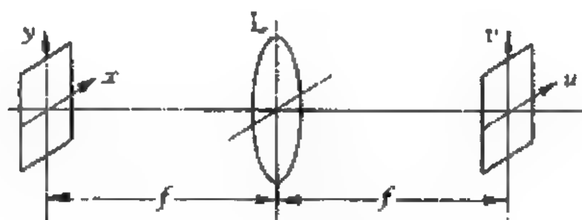


图 2.4 透镜傅里叶变换器

如果输入复振幅函数为 $f(x,y)$，则透镜后焦面的复振幅函数就是 $f(x,y)$的傅里叶变换，记为 $F(u,v)$，在这里略去了不重要的相位因子. 我们把 1.4 节(19)式重新改写如下：

$$F(u,v) = \iint_{-\infty}^{\infty} f(x,y)\exp\left[-\mathrm{i}\,\frac{2\pi}{\lambda f}(xu + yv)\right]\mathrm{d}x\mathrm{d}y, \qquad (1)$$

式中 λ 为波长，f 为透镜焦距. 从数学的角度来讲，$f(x,y)$也可以由 $F(u,v)$直接导出，同样略去不重要的相位因子，有

$$f(x,y) = \iint_{-\infty}^{\infty} F(u,v)\exp\left[\mathrm{i}\,\frac{2\pi}{\lambda f}(xu + yv)\right]\mathrm{d}u\mathrm{d}v, \qquad (2)$$

(1)式和(2)式分别为光学傅里叶变换及其逆变换. 傅里叶变换及其逆变换可以用 $\mathscr{F}\{\cdot\}$及 $\mathscr{F}^{-1}\{\cdot\}$表示，见节 1.1.

两个傅里叶变换器可以串联起来，用两个透镜 L_1 和 L_2 构成著名的 $4f$ 系统，如图 2.5 所示．在输入平面、谱平面（两个透镜的共同焦面）和输出平面（$\xi\eta$ 平面）上，分别是输入信号函数 $f(x,y)$、它的傅里叶谱 $F(u,v)$ 及输出信号函数 $f(\xi,\eta)$．正因为 $F(u,v)$ 是空间频率的函数，傅里叶变换平面又称频率平面，在本书中，几种称呼在不同的场合下都会用到．由于透镜只能实现傅里叶变换，为了与逆变换等效，我们把输出平面旋转 180°，坐标轴 ξ 和 η 分别对应于 $-x$ 和 $-y$.

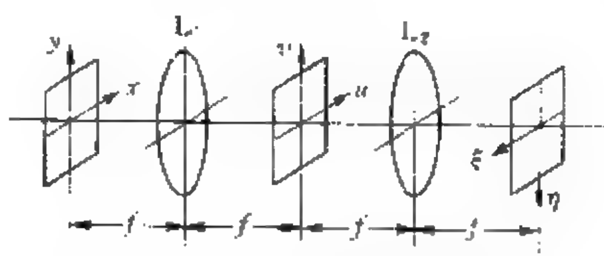


图 2.5　4f 光学信息处理系统
在输入平面、谱平面和输出平面上分别为
输入函数、傅里叶变换函数及输出函数

自从薄透镜（或透镜系统）对于相干光的复振幅信号的傅里叶变换特性被发现以后，光学信号的频谱就不再是一个抽象的数学概念，而变成了物理现实．在实用中，$f(x,y)$ 既可以是透明胶片或干版，又可以是空间光调制器（见本书第七章），用准直的相干光照射振幅透过率为 $f(x,y)$ 的光学模板，就产生了振幅函数 $f(x,y)$，在频率平面上出现它的傅里叶谱 $F(u,v)$．在频率平面上设置各种滤波器，可以对信号频谱进行改造，滤掉不需要的信息及噪声，提取或增强我们感兴趣的信息．经滤波的频谱通过第二个透镜还原成经改造的信号 $f(\xi,\eta)$（实际上它与原来的函数已有了不同程度的差别），呈现在输出平面上．注意所有的探测器，包括眼睛，都只能探测到光强，即振幅的模的平方．

借助于符号 $\mathscr{F}$，可以把(1)及(2)式表为

$$F(u,v)=\mathscr{F}\{f(x,y)\} \tag{3}$$

及

$$f(x',y') = \mathscr{F}\{F(u,v)\}. \tag{4}$$

这里(x',y')是输出平面上的坐标,坐标轴方向与(x,y)相同,它可以用傅里叶逆变换表示如下:

$$f(-x',-y') = \mathscr{F}^{-1}\{F(u,v)\}, \tag{5}$$

由图 2.5,显然有$\xi=-x'$及$\eta=-y'$,从而得到

$$f(\xi,\eta) = \mathscr{F}^{-1}\{F(u,v)\}. \tag{6}$$

这样一来,顺序进行的两次变换可以用图 2.6 表示.



图 2.6　包含傅里叶变换及逆变换的傅里叶处理系统

2.4　线性系统与卷积

如果用一个附加的滤波器与$F(u,v)$相乘,图 2.6 所示的系统确实表现为线性系统. 一个线性系统可定义如下:设

$$g_1(\xi,\eta) = \mathscr{L}\{f_1(x,y)\}, \tag{1}$$

$$g_2(\xi,\eta) = \mathscr{L}\{f_2(x,y)\}, \tag{2}$$

则有

$$\alpha g_1(\xi,\eta) + \beta g_2(\xi,\eta) = \mathscr{L}\{\alpha f_1(x,y) + \beta f_2(x,y)\}, \tag{3}$$

式中α,β为常数.

函数$f(x,y)$和$h(x,y)$的卷积我们已在节 1.1 中定义,现在重新写在下面:

$$\begin{aligned} f(x,y) * h(x,y) &= g(x,y) \\ &= \iint_{-\infty}^{\infty} f(\xi,\eta)h(x-\xi,y-\eta)\mathrm{d}\xi\mathrm{d}\eta. \end{aligned} \tag{4}$$

为了证明卷积是线性变换,引入

$$g_1(x,y) = \iint_{-\infty}^{\infty} f_1(\xi,\eta)h(x-\xi,y-\eta)\mathrm{d}\xi\mathrm{d}\eta, \tag{5a}$$

及

$$g_2(x,y)=\iint\limits_{-\infty}^{\infty} f_2(\xi,\eta)h(x-\xi,y-\eta)\mathrm{d}\xi\mathrm{d}\eta. \tag{5b}$$

我们立即得到

$$\alpha g_1(x,y)+\beta g_2(x,y)$$
$$=\iint\limits_{-\infty}^{\infty}[\alpha f_1(\xi,\eta)+\beta f_2(\xi,\eta)]h(x-\xi,y-\eta)\mathrm{d}\xi\mathrm{d}\eta. \tag{6}$$

这样就证明了卷积是线性运算.

如果输入函数是函数 $\delta(x,y)$,则输出

$$g(x,y)=h(x,y), \tag{7}$$

$h(x,y)$称为系统对脉冲的响应,简称脉冲响应. 当输入是一个点或一个脉冲时,其振幅是函数 $\delta(x,y)$, 输出振幅函数即$h(x,y)$,观察到的光强函数则为$|h(x,y)|^2$,它表示一个物点所形成的像的弥散,称点扩散函数.

Elias[12]第一个把电子学的概念引入光学之中. 如图 2.7 所示,成像过程可以看成是线性变换,图 2.7 表示一个输入的图像经过线性系统变成了输出图像. 该线性系统可以是成像透镜,也可以是光-电混合系统. 例如阴极射线管(CRT).



图 2.7 传递图像的线性系统

考虑由透镜产生的像. 由于透镜的限制,物平面上一个点的像可能有弥散,弥散的范围取决于透镜的质量及参数. 我们把原始的物体看成是大量点的集合,则该物体通过光学系统形成的像将是同样数量的弥散的光斑的集合,这些弥散斑由脉冲响应$h(x,y)$表示,脉冲响应是透镜的特征.

上面所述成像过程也适用于非相干情况,但 $f(x,y)$,$h(x,y)$

和 $g(\xi,\eta)$均为光强函数，此外，在非相干情况下，$h(x,y)$直接表示点扩散函数，不需要求平方.

2.5 空间滤波

运用同样的变量(x,y)，上一节(4)式可改写为

$$f(x,y)*h(x,y)=g(x,y)$$

$$=\iint_{-\infty}^{\infty}f(\alpha,\beta)h(x-\alpha,y-\beta)\mathrm{d}\alpha\mathrm{d}\beta. \quad (1)$$

设 f,h 和 g 的傅里叶变换分别为F,H 和G,则根据节 1.1 中关于卷积的变换定理,我们得到

$$G(u,v)=F(u,v)H(u,v). \quad (2)$$

设 $H(u,v)$为脉冲响应 $h(x,y)$的傅里叶变换，则 $H(u,v)$称为传递函数，它表征系统对输入信号的传递性能，使输入信号转换成输出信号. 一般来讲，$H(u,v)$使输入信号的频谱 $F(u,v)$转换成 $G(u,v)$，它是 $F(u,v)$和 $H(u,v)$的乘积. 于是，我们可以把图 2.7 所示的成像一类的线性系统的作用，等价为图 2.6 所示的两步的过程. 换句话说，成像这一类线性操作可以用图 2.6 所示的系统来模拟，或者用图 2.5 所示的光学系统来实现. 进一步，一个畸变像可以借助于图 2.5 所示的系统来校正. 这样改变频谱成分的操作称为空间频率滤波，简称空间滤波.

图 2.8 是一个空间滤波的例子. 图 2.8(a)是一个原始像，图像上叠加了具有光栅结构的高频噪声. 把这一个图像的透明片放在输入平面上，用准直的相干光照亮，它的频谱就出现在频率平面上，其中低频成分分布在频率平面的原点附近，而高频成分则远离原点. 用一个带小孔的不透明模板(又称针孔滤波器)，放在频率平面上，使小孔与原点一致，就构成低通滤波器，因为低频成分通过小孔，而高频成分被阻拦，输出图像不再带有高频成分，照片上就不出现光栅结构，见图 2.8(b). 反之，我们也可以通过拦去频

率平面中心的低频成分的方法加强高频信号，称高通滤波器. 图 2.9 是高通滤波的例子. 图 2.9(a)是原始图像，经高通滤波后的图像见图 2.9(b). 由于图像的边缘包含较丰富的高频成分，滤波的结果突出了边缘，在图像处理中称边缘增强.



图 2.8 (a) 输入图像 (b) 用针孔滤掉高频的输出图像

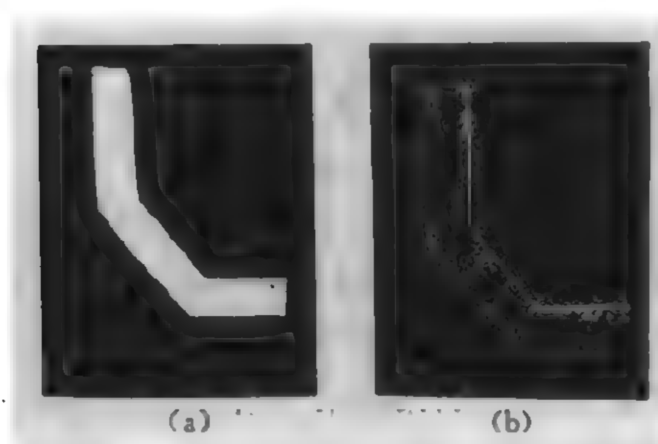


图 2.9 (a) 输入图像
(b) 用一个很小的不透光模板遮挡频率平面中心，滤去低频成分后的输出图像

2.6 照相图像的恢复

照片的复原可以说是空间滤波的第一个应用，这些工作是由 Marechal 和 Croce[8]，以及稍后的 Tsujiuchi[10] 完成的. 在上一节中我们曾讲过图像是点的集合，由于衍射效应，透镜的像是弥散光

斑的集合．如果透镜的像形成在理想成像位置上，它是清晰的，具有较高的反差．如果成像的位置偏离理想像面(又称“离焦”)，光斑就将更加弥散，使像变得更加模糊，在照相时，由于未准确对焦，经常会照出模糊的相片，就是例子．以上两篇文章的作者指出，离焦的模糊图像可以借助于空间滤波技术加以恢复．

设原始图像为 $f(x,y)$，由于离焦，像点变成了一个光斑，用 $h(x,y)$ 表示．结果，输出图像为输入图像和点扩散函数 $h(x,y)$ 的卷积：

$$g(x,y)=f(x,y)*h(x,y)$$
$$=\iint_{-\infty}^{\infty} f(\alpha,\beta)h(x-\alpha,y-\beta)\mathrm{d}\alpha\mathrm{d}\beta.$$

在初级近似下，$h(x,y)$ 表示为一个圆形的光斑，它的傅里叶变换见图 2.10(a)，我们略去了数学推导．由于它是旋转对称的，可表为 $h(r)$，见图 2.10(b)，$h(r)$ 的傅里叶变换为 $H(\rho)$，见图 2.10(c)．

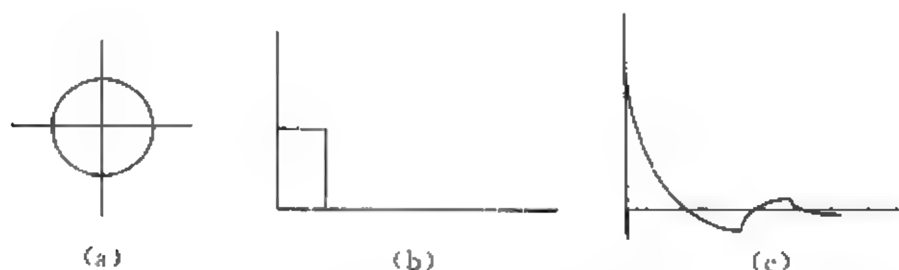


图 2.10 (a) $h(x,y)$ 表示一个圆形光斑 (b) 圆对称函数 $h(r)$ (c) $h(r)$ 的傅里叶变换 $H(\rho)$

如果成像平面就是理想像面，则 $h(x,y)$ 近似为 δ 函数，圆对称函数 $h(r)$ 也成为 δ 函数，其傅里叶变换为常数．

设理想平面上的像为 $f(x,y)$，其傅里叶变换则为 $F(u,v)$，离焦象为 $f(x,y)*h(x,y)$，其傅里叶变换则为 $F(u,v)H(u,v)$．为了获得清晰图像，频谱函数必须是 $F(u,v)$，那么，$F(u,v)H(u,v)$ 必须乘以 $1/H(u,v)$ 以产生 $F(u,v)$．我们首先构造一个振幅滤波

器和一个相位滤波器，参见图 2.11，振幅滤波器上的不同透过率可以用镀膜方法实现，相位滤波器则仅提供符号 +1 和 −1，使 $H(\rho)$ 与两个滤波器的连乘积为常数，则两个滤波器的乘积就表示 $1/H(\rho)$（差一个常数因子）. 当我们用这一复合的滤波器（称“逆滤波器”）对 $F(u,v)$ 进行滤波后，在离焦的像面上将生成清晰图像 $f(x,y)$. 图 2.12(a) 表示离焦像，图 2.12(b) 表示经过校正复原的像. 关于模糊图像恢复的进一步讨论，可参见杨振寰的著作[21].

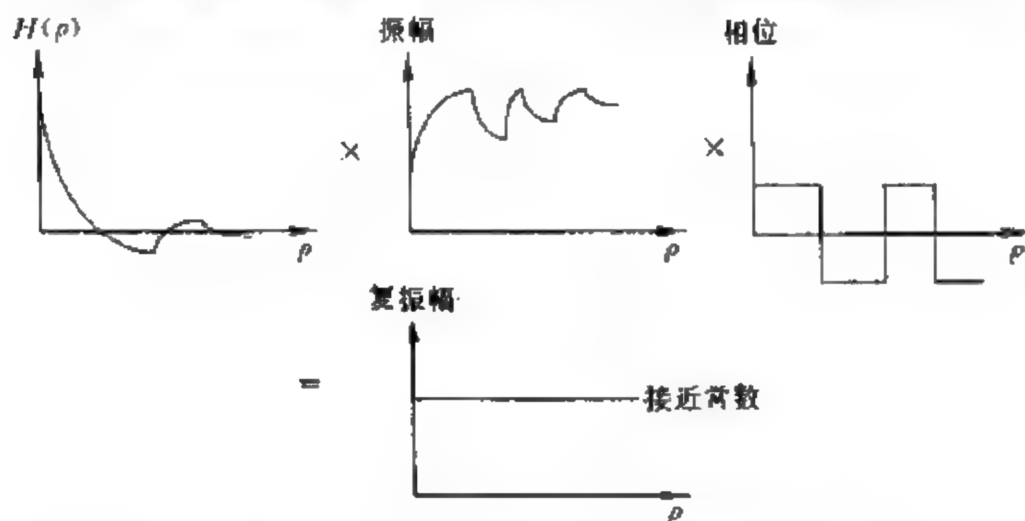
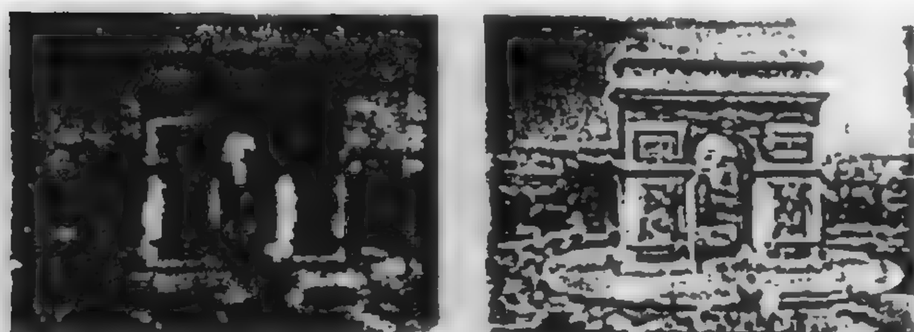


图 2.11 $H(\rho)\times$振幅滤波器$\times$相位滤波器 $\approx$ 常数



(a) 离焦像　　(b) 经过校正复原后的像[13]

图 2.12

2.7 全 息 术

全息术原来是盖柏(Gaber)在 1948 年发明的，它是一个记录

和重建波前的技术．设想我们在房间里，透过玻璃窗去看一个窗外的三维物体，例如一朵花．花透过玻璃窗最终在人眼的视网膜上成像．物理学仅仅承认近距作用，不承认超距作用．从近距作用的观点来看，物体的存在并不是看见物体的直接原因，只要由物体发生的光波（称为“物光”）射到人眼的视网膜上，人就能看到这个物体，看见物体的直接原因是物光，只要存在一个与物光完全相同的光波，即使物体不存在，或者墙上根本没有开窗，我们都能够“看见”这个物体——花．物体辐射（或反射）的光波的波前由振幅和相位构成，因此，只要在光波经过的路径的某一平面上设法记录并重建这个波前，从而完整地重建物光，我们就可以再次看到这朵花，不管它是否真的存在．全息术是通过记录物光与另一参考光的干涉图来记录波前的，下面就来讨论这一方法．

全息术原理可简单描述如下：设从一个三维物体（例如一朵花）上发出的复杂的物体光波在它通过的某一平面上的复振幅分布为 $O(u,v)$，它的振幅及相位分别为 $|O(u,v)|$ 及 $\phi(u,v)$，亦即

$$O(u,v) = |O(u,v)|\exp[\mathrm{i}\phi(u,v)]. \tag{1}$$

当我们在该平面上放置一个记录材料，例如照相干版，它记录了物光的光强分布，即

$$|O(u,v)|^2 = |O(u,v)|\exp[\mathrm{i}\phi(u,v)]\cdot|O(u,v)|\exp[-\mathrm{i}\phi(u,v)]. \tag{2}$$

显然在记录过程中丢失了相位信息，而且被丢失的相位信息无法恢复．当干版冲洗出来后，它记录的是 $|O(u,v)|^2$，当我们观察这个记录时，我们显然看不到原来的花．

现在让我们在记录过程中添加一个参考光 R，参考光与物光发生干涉，因而干版上记录的将是干涉图样，这一记录就称为全息图，它记录的光强分布为

$$\begin{aligned}&|O(u,v)+R(u,v)|^2\\&\quad=|O(u,v)|^2+|R(u,v)|^2\\&\qquad+O(u,v)R^*(u,v)+O^*(u,v)R(u,v),\end{aligned} \tag{3}$$

经冲洗后的全息图的透过率正比于记录的光强．当用原来记录全息图的参考光单独照射全息图时，透过全息图的光波复振幅为

$$
\begin{aligned}
&|O(u,v)+R(u,v)|^2R(u,v) \\
&\quad =|O(u,v)|^2R(u,v)+|R(u,v)|^2R(u,v) \\
&\qquad +O(u,v)|R(u,v)|^2+O^*(u,v)R(u,v)R(u,v). \qquad (4)
\end{aligned}
$$

如果 $R(u,v)$是一个均匀的光场，则上式第三项正比于物光的复振幅 $O(u,v)$，它正是原来从物体辐射的物体光波，当该光波射入我们的眼睛时，我们就重新看到了原来的三维物体——花，尽管此时物体已不存在，我们却从全息图中看到的这个"虚"物体和原物的形状、大小、位置相同，因而是立体的、逼真的、生动的，完全是真实物体的精确复制品．

值得注意的是(4)式中的第四项，它是物光的复共轭波．由于物光一般是发散的光波，对应的是"虚"物，就好比物体在镜中的像；物光的共轭波却是会聚的，它的焦点给出物体的"实"像．在本书第九章"互泵浦相位共轭"一节中，还将介绍利用光折变晶体的相位共轭产生全息实像的效应．

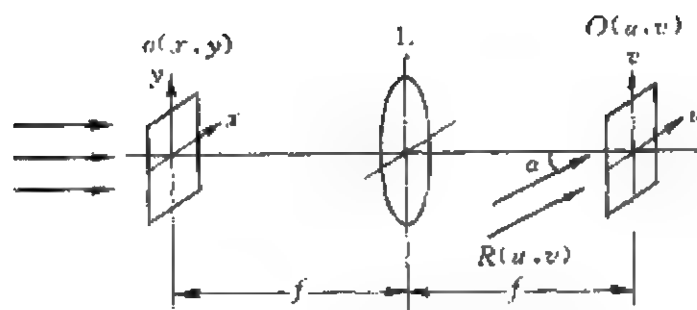


图 2.13 傅里叶变换全息图的记录

在信号处理的诸多应用中，输入信号 $o(x,y)$ 记录在透明片上，用准直的相干光照明，在透镜 L 的后焦面上出现 $o(x,y)$ 的傅里叶变换 $O(u,v)$，全息图记录 $O(u,v)$ 和一个倾斜的平面参考光波 $R(u,v)$ 的干涉图，R 的振幅为 1，可表为

$$R(u,v) = \exp[\mathrm{i}au], \tag{1}$$

其中

$$a = 2\pi\sin\alpha/\lambda, \tag{2}$$

α 是参考光与光轴的夹角，λ 为波长. 以(1)式代入节 2.7(3)式，得到

$$\begin{aligned}&|O(u,v) + R(u,v)|^2\\&\quad = |O(u,v)|^2 + 1 + O(u,v)\exp[-\mathrm{i}au]\\&\qquad + O^*(u,v)\exp[\mathrm{i}au].\end{aligned} \tag{3}$$

当用参考光再现物光时，透过全息图的光波波前为

$$\begin{aligned}&|O(u,v) + R(u,v)|^2 R(u,v)\\&\quad = (|O(u,v)|^2 + 1)\exp[\mathrm{i}au]\\&\qquad + O(u,v) + O^*(u,v)\exp[\mathrm{i}2au],\end{aligned} \tag{4}$$

式中等号右边第一项仍沿着参考光的方向传播，与光轴夹角为 α；第二项是重建的物光，沿光轴传播；第三项则沿着与光轴的夹角为 2α 的方向传播，见图 2.14.

全息图也可以通过图 2.15 所示的傅里叶变换器再现. 在这个系统中，重建全息图的参考光与光轴平行. 重建的波前直接由(3)式表示，它的傅里叶逆变换由下式给出：

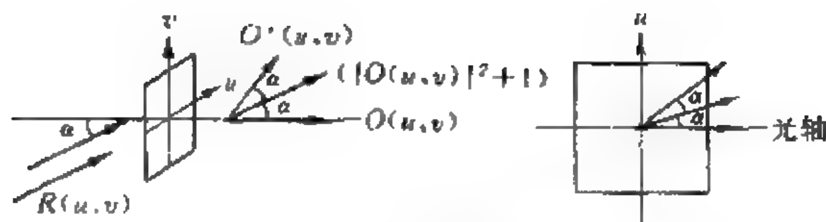


图 2.14 傅里叶变换全息的再现

$$
\begin{aligned}
&\mathscr{F}^{-1}\{|O(u,v))|^2+1+O(u,v)\exp[-\mathrm{i}au]+O^*(u,v)\exp[\mathrm{i}au]\}\\
&=\mathscr{F}^{-1}\{|O(u,v)|^2+1\}+o(\xi-b,\eta)+o^*(-\xi+b,-\eta),
\end{aligned}
\tag{5}
$$

在推导中我们用到节 1.1 中的宗量平移的傅里叶变换性质，并注意透镜傅里叶变换的频率变量为 $u/\lambda f$. 这样一来，原来的物体 $o(\xi,\eta)$ 在输出平面上 $(b,0)$ 处出现，它的复共轭镜面对称像 $o^*(-\xi,-\eta)$ 则在 $(-b,0)$ 处出现，

$$
b=f\sin\alpha, \tag{6}
$$

而 $\mathscr{F}^{-1}\{|O(\xi,\eta)|^2+1\}$ 项则位于输出平面中心.

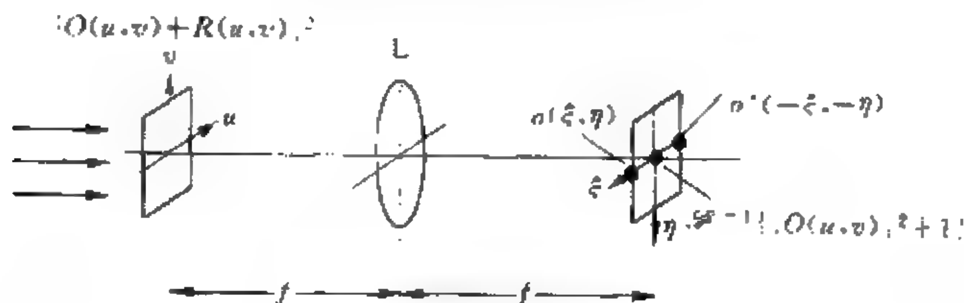


图 2.15 用傅里叶变换器进行傅里叶变换全息的再现

傅里叶变换全息的应用之一，是以全息图的形式进行光学存储. 全息存储的基本概念是在 60 年代，在 Van Heerden 的有创见的论文中提出来的[15,16]. 直接的图像存储，例如常规高密度缩微胶片，片基上的一点点疵病都会引起部分记录的丢失，而且丢失的信息是永远无法恢复的. 然而，当使用全息图进行高密度存储时，在片基上的疵病诸如划痕和灰尘都不会破坏信息，只是在信息再

现时稍微增加了一点噪声，因而瑕疵并不意味着某一部分信息的损失. 设想被存储的信息是一串二进制代码，则该代码以二维的形式存储在全息图中，称为"页存储". 以傅里叶变换全息图进行页存储的最大优点在于它对应于最小的空间带宽积，且任意一个代码都存储在整个全息图平面上. 我们已经讨论了记录从花朵上辐射光波的全息图，如果把它比喻为一扇窗，通过这扇窗，我们不仅看到了花，还可以看到许多其他物体. 这扇窗可以透过无限个物体发出物光的波前，因为这些光束是独立传播的，不会发生相互干扰. 然而实际上全息图只能记录有限个波前，因为波前是以干涉图的形式进行记录的，所能记录的波前的数目取决于全息图的分辨率和动态范围. 显然，全息图越大，存储的容量就越大.

常规成像把图形存储在一个平面中，然而全息技术却把二维图像存储在记录介质的三维干涉图中[17]，在全息再现时，记录下来的三维干涉图又还原成二维图形. 三维干涉图称体积全息，简称体全息. 以三维记录介质进行存储使存储容量显著加大.

一个有趣的事实是，在常规全息术中，一张二维的平面全息记录并显示一个三维的物体，例如花朵，然而在数字光存储技术中，三维的体全息记录并显示的却是二维的页面信息.

全息存储的另一优点在于多个全息图可以存储在一张全息图中. 二进制码通常不能进行多重存储，因为我们无法将它们分离. 然而，采用不同的记录角度，多重全息图可以分别再现. 90 年代初，Mok 等成功演示了在一个 2 cm×1.5 cm×1 cm 的掺铁铌酸锂晶体中存储的 5000 个全息图的实验[19].

2.9 相关和卷积

全息图的主要优点在于它可以记录和重建波前，因此可以用全息图来产生一个复数函数滤波器，而不必像图 2.11 那样采用相位滤波器和振幅滤波器进行复合滤波. 全息术的重要应用之一，

是实现 Vander Lugt[20]提出的复数匹配滤波，该复数滤波器就是图 2.13 所示的傅里叶变换全息图. 输入函数是 $g(x,y)$而不是 $o(x,y)$. 由节 2.8(3)式，全息图的透过率分布函数为

$$\begin{aligned} T(u,v) &= |G(u,v)+R(u,v)|^2 \\ &= |G(u,v)|^2 + 1 + G(u,v)\exp[-\mathrm{i}au] \\ &\quad + G^*(u,v)\exp[\mathrm{i}au]. \end{aligned} \tag{1}$$

该全息图又称复匹配滤波器，再放回图 2.16 所示系统的傅里叶变换平面上. 把一个新的输入 $f(x,y)$放到输入平面上去，在傅里叶平面上产生它的谱函数 $F(u,v)$，在紧靠复数滤波器后面的场为

$$\begin{aligned} F(u,v)T(u,v) &= (|G(u,v)|^2+1)F(u,v) \\ &\quad + F(u,v)G(u,v)\exp[-\mathrm{i}au] \\ &\quad + F(u,v)G^*(u,v)\exp[\mathrm{i}au]. \end{aligned} \tag{2}$$

$4f$ 系统的第二个透镜在输出平面产生 $F(u,v)T(u,v)$的傅里叶变换. 与节 2.8(5)式类似，第一项的傅里叶变换位于输出平面的中心，$F(u,v)G(u,v)$和 $F(u,v)G^*(u,v)$对应的变换项则出现在输出平面$(b,0)$及$(-b,0)$处.

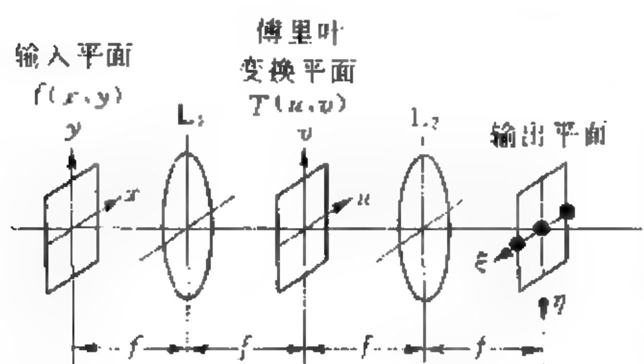


图 2.16　复匹配滤波

现在我们可以来分析 $F(u,v)G(u,v)$ 及 $F(u,v)G^*(u,v)$ 的傅里叶逆变换. 节 1.1 关于卷积和相关的变换法则表明，$F(u,v)G(u,v)$的逆变换就是 $f(\xi,\eta)$和 $g(\xi,\eta)$的卷积，$F(u,v)G^*(u,v)$经傅里叶逆变换后则生成 $f(\xi,\eta)$和$g(\xi,\eta)$的相关，众所周知，当 $f(x,y)$和$g(x,y)$相同对，相关输出达到极大. 参见图 2.17.

由图 2.17 可知，相关可以作为两个图形相似性的度量，如果待识别图形 $f(x,y)$和目标 $g(x,y)$全同，相关输出 $o(\xi,\eta)$将包含一个明显的相关峰；如果 $f(x,y)$和 $g(x,y)$不同，相关输出 $o(\xi,\eta)$

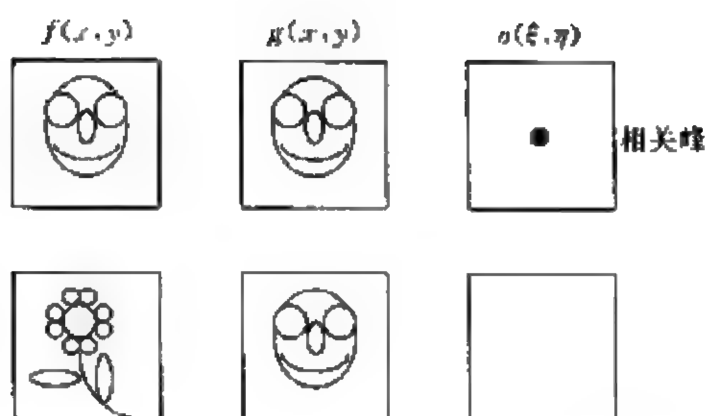
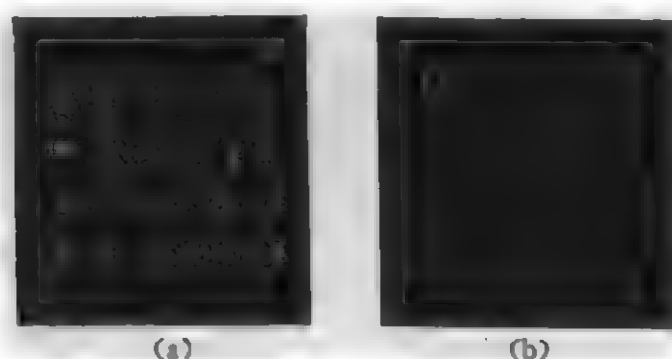


图 2.17　相关的图示

将是平坦的，具有很小的起伏. 因此，光学相干器广泛地用于自动图形识别. 被识别的图形记录在复匹配滤波器中，滤波器放置于傅里叶平面上. 当输入信号就是 $g(x,y)$时，输出平面上出现尖锐的相关峰. 可以用CCD来扫描相关输出平面加以探测. 相关峰的出现表明输入图形与目标相同或相似，相关峰的信噪比则表征相似的程度. Vander Lugt 当时的实验结果显示在图 2.18 之中. 我们将在第四章中详细讨论光学图形识别.



(a)　　(b)

图 2.18　(a) 输入物体　(b) 相关峰指示了字母"g"[20]

2.10　结　　论

本章在相干光透镜傅里叶变换的基础上介绍了光学信息处理的基本概念. 我们列举了若干早期和近代的应用的例子，说明相

干光信息处理可以实现空间滤波、使相位物体成为可见、全息成像、离焦照片的恢复、光学信息的存储和再现，及应用全息匹配滤波器实现卷积和相关.

参考文献

[1] F. T. S. Yu, Optics and Information Theory, R. E. Krieger, Malabar, Florida, 1984.

[2] F. T. S. Yu and S. Jutamulia, Optical Signal Processing, Computing, and Neural Networks, Wiley, New York, 1992.

[3] E. Abbe, Beitrage zur Theorie des Mikroskops und der mikroskopichen Wahrnehmung, *Archiv. Mikroskop. Anat.*, **9** (1873), 413-468.

[4] A. B. Porter, On the diffraction theory of microscopic vision, *Phil. Mag.*, **11** (1906), 154-166.

[5] F. Zernike, Das Phasenkontrastverfahren bei der mikroskopischen Beobachtung, *Z. Tech. Phys.*, **16**(1935), 454-456.

[6] See for example, E. Hecht and A. Zajac, Optics, Addison-Wesley, Reading, Massachusetts, 1974, p. 480.

[7] P. M. Duffieux, L'Integrale de Fourier et ses Applications a l'Optique, Faculte des Sciences, Besancon, 1946.

[8] A. Marechal and P. Croce, Un filtre de frequences spatiales pour l'amelioration du contraste des images optiques, *Compt. Rend. Acad. Sci. Paris*, **237**(1953), 607-609.

[9] H. H. Hopkins, On the diffraction theory of optical images, *Proc. Roy. Soc.*, **A 217**(1953), 408-432.

[10] E. L. O'Neil, Spatial filtering in optics, *IRE Trans. Inf. Theory*, **IT-2** (1956), 56-65.

[11] L. J. Cutrona, E. N. Leith, C. J. Palermo, L. J. Porcello, Optical data processing and filtering systems, IRE Trans. Inf. Theory, **IT-6**(1960), 386-400.

[12] P. Elias, D. S. Grey, D. Z. Robinson, Fourier treatment of optical processes, *J. Opt. Soc. Am.*, **42**(1952), 127-134.

[13] J. Tsujiuchi, Correction of optical images by compensation of aberrations and by spatial frequency filtering, *Progress in Optics*, **2**(1963), 133-180.

[14] D. Gabor, A new microscope principle, Nature, **161**(1948), 777-778.

[15] E. N. Leith and J. Upatnieks, Reconstructed wavefronts and communication theory, *J. Opt. Soc. Am.*, **52**(1962), 1123-1130.

[16] P. J. van Heerden, A new optical method of storing and retrieving information, *Appl. Opt.*, **2**(1963), 387-392.

[17] P. J. van Heerden, Theory of optical information storage in solids, *Appl. Opt.*, **2** (1963), 393-400.

[18] S. Jutamulia and G. M. Storti, Three-dimensional optical storage, in Optical Storage and Retrieval, F. T. S. Yu and S. Jutamulia, Eds., Marcel Dekker, New York, 1996.

[19] F. H. Mok, Angle-multiplexed storage of 5000 holograms in lithium niobate, *Opt. Lett.*, **18**(1993), 915-917.

[20] A. Vander Lugt, Signal detection by complex spatial filtering, *IEEE Trans. Inf. Theory*, **IT-10**(1964), 139-145.

[21] 杨振寰,光学信息处理,(母国光,羊国光,庄松林译),南开大学出版社,1986.

第三章　非相干光学信息处理

由于廉价的激光器的广泛应用，非相干光学信息处理已变得不那么重要了，与相干光学信息处理相比，非相干光学信息处理的优势很少．现在很少有人试图去建立一个非相干光学信息处理器，例如非相干光学相关器（参见本书第四章）．尽管如此，大量光学仪器仍是采用非相干光或自然光作为光源的，其中大多数是成像仪器，例如照相机、显微镜、望远镜、投影仪、制版设备等等．应当说，常规意义下的成像，也是光学信息处理的重要应用，在这个意义上，非相干光学信息处理的基本概念仍然有必要加以研究，这些概念已广泛应用于非相干光成像之中．

3.1　杨氏干涉仪和空间相干性

干涉仪是产生光波干涉的仪器或装置，仅仅相干光才能产生干涉效应．因此干涉仪是研究光的干涉性的恰当的设备．常见的杨氏干涉仪是由双狭缝或双孔构成的．图 3.1 为杨氏干涉仪示意图．

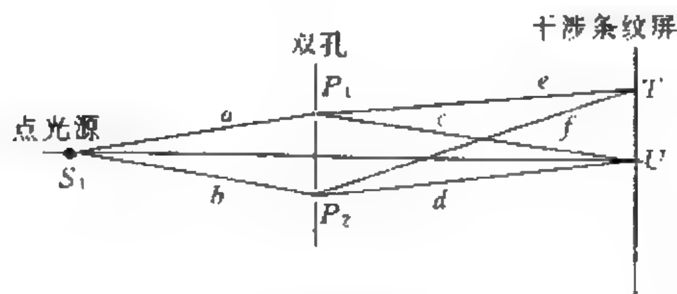


图 3.1　同轴点光源杨氏干涉仪

光轴上一点 S_1（例如一个原子）辐射的光波透过不透光屏上

的小孔 P_1，沿光路 $a-c$ 照亮图像屏上的一点 U. 另一束光波透过小孔 P_2，沿光路 $b-d$ 照亮 U. 由于两个光波是同一点光源辐射的，它们在 U 点发生干涉. 由于 $(a+c)=(b+d)$，U 为亮点. 若在 T 点满足条件 $(b+f)=(a+e+\lambda)$，λ 为光波波长，则 T 点也是亮点. U 和 T 中间的点是暗点，亦即干涉相消的点出现在光程差为 $\lambda/2, 3\lambda/2, 5\lambda/2, \cdots$ 之处. 干涉加强的点出现在光程差为 $\lambda, 2\lambda, 3\lambda, \cdots$ 之处. 这样一来在图像屏上将出现明暗相间的条纹. 如果在屏上能得到相干的条纹的话，就说照明小孔的光波是相干的.

如果点光源位于轴外，则干涉条纹也将发生位移，亮纹将在 W，V 等处出现，如图 3.2 所示. 此时仍然有 $(a+c)=(b+d)$，和 $(b+f)=(a+e+\lambda)$.

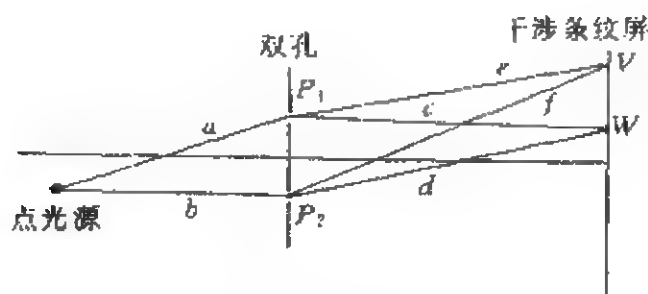


图 3.2　离轴点光源杨氏干涉仪

如果仅有一个点光源，例如 S_1 或 S_2，则我们可以观察到高对比度的相干条纹；如果 S_1 和 S_2 同时存在，我们将看到两个独立的干涉图样的非相干叠加，因为 S_1 和 S_2 是非相干的.

我们还可以考虑光源在一定的时间间隔内辐射的“波列”，它可以表为余弦函数 $\cos(2\pi x/\lambda - 2\pi\nu t + \varphi)$，式中 λ 为光波波长，ν 为光波频率，φ 为初相位. 频率、相位由光源决定，而波长则取决于光在介质中传播的速度，波长等于介质中的光速除以频率. 由于一个波列中的 ν 和 φ 均为常数，因而上述光扰动作为 x 或 t 的函数是有规律的. 光波可以在空间加以分割并重新合成，以产生干涉图样. 如果重新合成的两束光的相位差为 $0, 2\pi, 4\pi, \cdots$，则出现干涉加

强，如果相位差为 $\pi, 3\pi, 5\pi, \cdots$，则出现干涉相消.

然而我们知道，自然界和人工制造的光源辐射的光波的波列都只有有限的长度. 如果干涉仪中两个光束来自同一光源先后辐射的两个不同的波列，或来自两个不同的光源，则干涉图样不出现. 因为两个不同波列或两个不同光源的频率和相位不相同，重新合成的两个光波的相位变化很快，对应的干涉图样变化也很快，没有哪一个探测器能探测到这些瞬时的干涉图，因而观察到的只是均匀的亮度.

这样一来，两个非相干图样的叠加产生的效果是它们的强度叠加，在图 3.1 中和 3.2 中若两个点光源 S_1 和 S_2 同时存在，我们将观察到两个干涉图样的强度叠加，屏上的 T, U, V 和 W 都是亮的，干涉条纹的反差将下降. 如果有一大群点源 $S_1, S_2, S_3, \cdots$ 则屏上所有的点都是亮的，干涉条纹完全消失，我们说照明光是非相干的.

因此杨氏干涉仪可以用来研究光波的相干性. 我们通过 P_1 和 P_2 两个小孔是否在屏上产生干涉条纹来确定照明这两点的光波是否相干. 如果屏上出现高反差的相干条纹，光波就是相干的；如果屏上出现均匀的照明，光波就是非相干的；如果屏上出现低反差的条纹，则照明光就是部分相干的. 以 P_1, P_2 的位置为函数的相干性表征光波在 P_1, P_2 的相干的程度，称为空间相干性.

为了进行空间相干性的测量，我们可以改变 P_1 和 P_2 的间距. 间距增大时，发生两个效应，一个是条纹间距的变小，另一个是条纹反差度的下降. 条纹反差度决定了空间相干性. 如果小孔的间距大于某一极限后屏上的条纹不再出现，则称此极限间距为空间相干宽度(spatial coherence width). 在相干光处理系统中，我们总是假定空间相干宽度大于光学系统的横向特征尺度，而在非相干光处理系统中，我们总是假定空间相干宽度为零. 而在部分相干光处理系统中，假定空间相干宽度大于零，并小于系统的特征尺度.

光的干涉效应的进一步定量分析，请参见文献[1]和[2].

3.2 非相干像的形成

在第二章中，我们讨论了相干光的成像过程. 相干光成像系统是线性系统，设在输入平面上有一个理想的点光源，表为 $\delta(x,y)$，它在输出平面上的像即系统的脉冲响应，表为 $h(x,y)$，相应的强度分布为 $|h(x,y)|^2$.

我们把输入的二维物体看成是大量点源的连续分布，在相干光的照明下，输出平面的复振幅分布就是所有的点源对应的脉冲响应光斑 $h(x,y)$ 的叠加. 如果输入物体的复振幅分布为 $f(x,y)$，输出像的复振幅分布即为

$$g(\xi,\eta) = \iint_{-\infty}^{\infty} f(x,y)h(\xi - x,\eta - y)\mathrm{d}x\mathrm{d}y, \tag{1}$$

即输出信号是输入信号与脉冲响应的卷积. 输出的光强分布则为

$$|g(\xi,\eta)|^2 = \left|\iint_{-\infty}^{\infty} f(x,y)h(\xi - x,\eta - y)\mathrm{d}x\mathrm{d}y\right|^2. \tag{2}$$

(1) 式在频域中的表达式为

$$G(u,v) = F(u,v)H(u,v), \tag{3}$$

其中 G,F 和 H 分别是 g,f 和 h 的傅里叶变换，$H(u,v)$ 又称成像系统的传递函数. 为了区别于非相干光成像系统的传递函数，通常称 $H(u,v)$ 为相干传递函数，简写为 CTF (coherent transfer function).

让我们考虑同一系统，仅将照明光改为非相干光. 复振幅的脉冲响应仍是 $h(x,y)$，相应的强度分布为 $|h(x,y)|^2$. 如前所述，设物体是大量点光源的连续分布，则该物体的像为所有点光源的像的叠加. 然而由于照明光为非相干光，从各个点光源辐射的光波彼此是不相干的. 与上一节中关于杨氏干涉仪的讨论相似，各点光源的像也是彼此不相干的，输出像是输入平面物体上各点的像的强度叠加，因此输出的复振幅分布无法计算，而强度分布则为

$$|g(x,y)|^2 = \iint_{-\infty}^{\infty} |f(x,y)|^2 |h(\xi - x, \eta - y)|^2 \mathrm{d}x\mathrm{d}y. \quad (4)$$

在频域中，上式变成

$$GI(u,v) = FI(u,v)HI(u,v), \quad (5)$$

这里我们用 GI，FI 和 HI 分别表示 $|g|^2$，$|f|^2$ 和 $|h|^2$ 的傅里叶变换. 光强的脉冲响应 $|h(x,y)|^2$ 又称点扩散函数，记为 PSF(point spread function)，而 $HI(u,v)$ 则称为非相干成像系统的传递函数，简称光学传递函数，简写为 OTF(optical transfer function).

由于 $H(u,v)$ 是 $h(x,y)$ 的傅里叶变换，根据傅里叶变换的法则(参见节 1. 1)，$|h(x,y)|^2$ 的傅里叶变换为 $H(u,v)$ 的自相关，亦即

$$HI(u,v) = \iint_{-\infty}^{\infty} H^*(p,q)H(p+u,q+v)\mathrm{d}p\mathrm{d}q. \quad (6)$$

上式表明 OTF 是 CTF 的自相关. OTF 通常是复函数，可表为

$$\mathrm{OTF} = |\mathrm{OTF}|\mathrm{e}^{i\phi}, \quad (7)$$

我们进一步记

$$\mathrm{MTF} = |\mathrm{OTF}|, \quad (8)$$

MTF 称为调制传递函数(modulation transfer function)，而相位函数 ϕ 则记为

$$\mathrm{PTF} = \phi, \quad (9)$$

称相位传递函数(phase transfer function)，有关光学传递函数的深入研究请参见文献[3].

3. 3 MTF 的测量

非相干成像系统的 MTF 可以借助于输入平面上的余弦光栅来测量，余弦光栅的光强分布为

$$i(x) = 1 + \cos(2\pi p_0 x), \quad (1)$$

设系统的输出为

$$o(x) = 1 + m\cos(2\pi p_0 x + \varphi), \tag{2}$$

式中反差度即调制度 m 可如下测出

$$m = \frac{o(x)_{\max} - o(x)_{\min}}{o(x)_{\max} + o(x)_{\min}}. \tag{3}$$

在频域中，输入函数可表为

$$I(p) = \delta(p) + \delta(p - p_0)/2 + \delta(p + p_0)/2. \tag{4}$$

根据上一节的讨论，输出信号可写作

$$o(p) = I(p)\mathrm{OTF}(p). \tag{5}$$

以(4)式代入(5)式，得到

$$o(p) = \mathrm{OTF}(0)\delta(p) + \mathrm{OTF}(p_0)\delta(p - p_0)/2 + \mathrm{OTF}(-p_0)\delta(p + p_0)/2. \tag{6}$$

通常的归一化手续规定

$$\mathrm{OTF}(0) = 1, \tag{7}$$

由于 OTF 是自相关函数，它必然具有对称性，所以有

$$\mathrm{OTF}(-p_0) = \mathrm{OTF}(p_0). \tag{8}$$

以(7)、(8)式代入(6)，得到

$$o(p) = \delta(p) + \mathrm{OTF}(p_0)[\delta(p - p_0) + \delta(p + p_0)]/2. \tag{9}$$

OTF 可借助于 MTF 和 PTF(用 ϕ 表示) 表为

$$\mathrm{OTF}(p_0) = \mathrm{MTF}(p_0)e^{i\phi}, \tag{10}$$

代入(9)式得到

$$o(p) = \delta(p) + \mathrm{MTF}(p_0)e^{i\phi}[\delta(p - p_0) + \delta(p + p_0)]/2, \tag{11}$$

上式的傅里叶逆变换为

$$o(x) = 1 + \mathrm{MTF}(p_0)\cos(2\pi p_0 x + \phi). \tag{12}$$

将上式与(2)式相比，得到

$$\mathrm{MTF}(p_0) = m, \tag{13}$$

而 PTF 则为

$$\phi = \varphi. \tag{14}$$

空间频率为 p_0 的调制传递函数 MTF 可根据公式(3)测得. 为

了获得完整的调制传递函数曲线，应对不同频率 p 的余弦光栅重复上述测量过程.

3.4 非相干空间滤波

在相干光学信息处理系统($4f$ 系统)中，第一透镜对输入图像进行傅里叶变换，第二个进行傅里叶逆变换. 第一透镜的后焦面即傅里叶变换域，又称傅里叶平面或频域. 因此物体的傅里叶变换可以从傅里叶平面上看到. 倘若物体是光栅，则我们在傅里叶平面上可以看到各个谱项. 然而，当我们把相干光源(激光)换成非相干光源(钨丝灯)，傅里叶平面上的傅里叶变换图像就消失了. 这一情形与杨氏干涉仪类似，当使用非相干光照相时杨氏条纹就消失了. 这是否意味着我们不能实现空间滤波?答案是否定的，尽管傅里叶变换看不到，我们仍然可以实现空间滤波.

设想我们在傅里叶平面上设置一个滤波器，用 $H(u,v)$ 表示. 为了方便起见，设该滤波器是一个简单的小窗口，表为 $H(u)$，系统的 CTF 显然就是 $H(u)$ 本身，而 OTF 则是CTF 的自相关，见图 3.3.

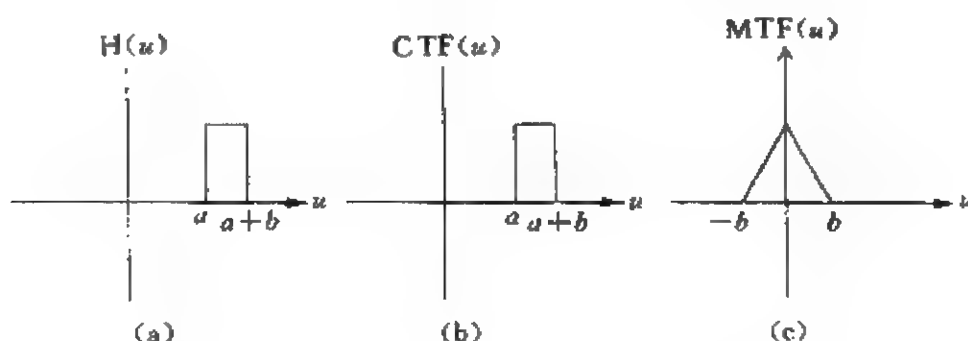


图 3.3　滤波平面上的实窗口函数生成的 CTF 及 OTF

为简单起见，图 3.3 只给出一维的情形. 由该图可以看出，尽管 CTF 是高通滤波器，通常为 $u = a$ 到 $u = a + b$，但 MTF 仍然

是低通滤波器，通常从 $u=-b$ 到 $u=b$，与 a 无关，表明非相干处理器不强调高频分量，它一般都是低通处理器. 有趣的事实在于，OTF 竟然与滤波器的位置 a 无关，仅取决于滤波器的宽度 b. 于是，由一组无规则分布的小孔构成的孔径的作用相当于低通滤波器. 这样一个滤波器的截止频率可以由针孔的直径导出，相当于图 3.3 中的 b. 使用这类滤波器的手续相当简单，如果用照相机去拍摄一个场景，该滤波器可以直接加在镜头上，拍得的照片中即不包含高频分量. 日常生活的经验告诉我们：当我们缩小照相机的光圈时，拍得的照片的分辨率（也就是"解析度"）下降，但景深加大.

Lohmann[4] 指出，Vander Lugt 相关器[5] 也能用在非相干光的情形. 相干 Vander Lugt 相关器的输出中，相关项为（参见节 4.3(14) 式）

$$c(\xi,\eta)=\iint_{-\infty}^{\infty}f(\alpha,\beta)g(\alpha-(\xi-b),\beta-\eta)\mathrm{d}\alpha\mathrm{d}\beta, \qquad (1)$$

其强度分布为

$$|c(\xi,\eta)|^2=\left|\iint_{-\infty}^{\infty}f(\alpha,\beta)g(\alpha-(\xi-b),\beta-\eta)\mathrm{d}\alpha\mathrm{d}\beta\right|^2. \qquad (2)$$

然而，当输入物体用非相干光而不是相干光照相时，相关项的强度分布变成

$$|c(\xi,\eta)|^2=\iint_{-\infty}^{\infty}|f(\alpha,\beta)|^2|g(\alpha-(\xi-b),\beta-\eta)|^2\mathrm{d}\alpha\mathrm{d}\beta, \qquad (3)$$

即 $|f|^2$ 和 $|g|^2$ 的相关. 因而当 f 与 g 全同时相关峰出现在 $(b,0)$ 处，也就是相干光处理器的相关峰位置. 然而在非相干情形下联合傅里叶变换器（JTC，参见节 4.8）不起作用. 联合傅里叶变换器实际上相当于杨氏干涉仪，而且两个小孔（或两个狭缝）的距离大于输入图形的横向尺寸. 根据节 3.1 的讨论可知，非相干情形下是看不到相干条纹的，因为非相干光的横向宽度几乎为零.

图 3.4 中给出非相干 Vander Lugt 相关器的实验结果. 在图中左下方没有相关信号，而右下方可以看到相关斑.

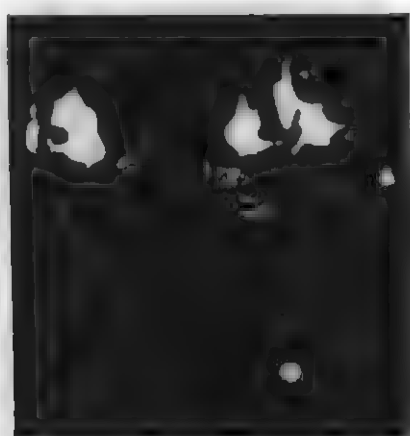


图 3.4　非相干 Vander Lugt 相关结果

3.5　迈克耳孙干涉仪和时间相干性

迈克耳孙干涉仪见图 3.5. 在迈克耳孙干涉仪中，输入的平行光被分光镜 BS 分成两束，分别被反射镜 M_1 和 M_2 反射，并通过分光镜重新合成，照到屏 S 上. 当两臂长度相等时($a = b$)，相干条纹出现. 注意两个反射镜应稍微倾斜一点，否则在屏上看不到条纹.

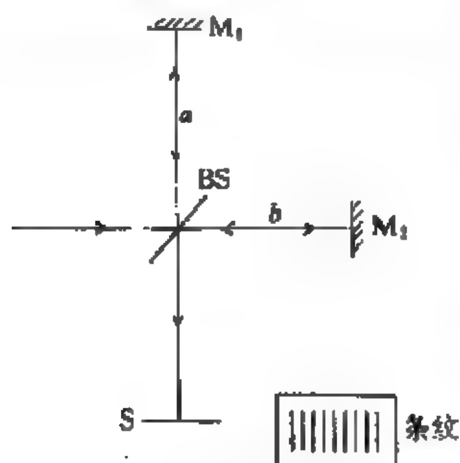


图 3.5　迈克耳孙干涉仪

若使得反射镜 M_2 沿光轴方向移远，使 $b > a$，干涉条纹的反差就会下降. 当 $(2b - 2a)$ 大于一定长度 l 后，屏上的条纹消失，变成均匀的亮斑，l 称光波的相干长度. 相干时间定义为

$$\tau = l/c, \tag{1}$$

式中 c 为光速. 在迈克耳孙干涉仪中，两个光束能够形成干涉条纹的前提条件是它们到达屏的时间差不大于 τ，或它们的光程差不大于 l；否则就不会产生干涉条纹.

相干长度 l 相当于节 3.1 中所谈到波列的平均长度. 因此对于一个给定的时刻，沿光波传播方向相干性度量体现为时间相干性，

在垂直于传播方向的截面中相干性度量体现为空间相干性.

3.6 傅里叶变换光谱仪

傅里叶变换不仅可用于光学信息处理,利用干涉仪,还可以用于光谱仪中,本节就来讨论这一课题.

让我们来考虑图 3.5 中屏上的一个点,称为观察点.该点的强度取决于两个相干光束的相位差.而相位差又取决于两光路的光程差 p. 由图 3.5,有

$$p = 2(b - a), \tag{1}$$

相位差还取决于波长 λ,

$$\lambda = c/\nu, \tag{2}$$

其中 c 为光速,ν 为频率.相干叠加的光强度为

$$I(p,\nu) = S(\nu)[1 + \cos(2\pi p\nu/c)], \tag{3}$$

式中 $S(\nu)$ 是产生干涉前的光强,称初始光强. $S(\nu)$ 表征了光波中的频率成分含量,正是我们感兴趣的光谱函数.当光程差为 p 时,在观察点探测到的总光强为

$$\begin{aligned} I(p) &= \int_0^\infty I(p,\nu)\mathrm{d}\nu, \\ &= \int_0^\infty S(\nu)\mathrm{d}\nu + \int_0^\infty S(\nu)\cos\left(\frac{2\pi p\nu}{c}\right)\mathrm{d}\nu. \end{aligned} \tag{4}$$

设

$$\int_0^\infty S(\nu)\mathrm{d}\nu = A, \tag{5}$$

它是与光程差无关的一个常量,则(4)式化作

$$I(p) = A + \int_0^\infty S(\nu)\cos\left(\frac{2\pi p\nu}{c}\right)\mathrm{d}\nu, \tag{6}$$

或

$$\int_0^\infty S(\nu)\cos\left(\frac{2\pi p\nu}{c}\right)\mathrm{d}\nu = I(p) - A. \tag{7}$$

我们一面移动第二块反射镜 M_2,一面在观察点测 $I(p)$,测得

足够稠密的 $I(p)$ 值.

$S(\nu)$ 的傅里叶逆变换为

$$s(x) = \int_{-\infty}^{\infty} S(\nu)\exp(-\mathrm{i}2\pi\nu x)\mathrm{d}\nu, \tag{8}$$

$S(\nu)$ 则可以用 $s(x)$ 表为

$$S(\nu) = \int_{-\infty}^{\infty} s(x)\exp(\mathrm{i}2\pi\nu x)\mathrm{d}x. \tag{9}$$

由于负的空间频率物理上不存在,它也不携带任何新的信息,在(8) 式中直接假定 $S(-\nu) = S(\nu)$,得到

$$s(x) = 2\int_{0}^{\infty} S(\nu)\cos(2\pi\nu x)\mathrm{d}\nu. \tag{10}$$

设

$$x = p/c, \tag{11}$$

代入(10) 式得到

$$s\left(\frac{p}{c}\right) = 2\int_{0}^{\infty} S(\nu)\cos\left(\frac{2\pi p\nu}{c}\right)\mathrm{d}\nu. \tag{12}$$

将上式代入(7) 式,得到

$$s(p/c) = 2[I(p) - A], \tag{13}$$

这里 $I(p)$ 和 A 都是可测量. 将(11) 式代入(9) 式得到

$$S(\nu) = \int_{-\infty}^{\infty} s\left(\frac{p}{c}\right)\exp\left(\frac{\mathrm{i}2\pi p\nu}{c}\right)\mathrm{d}\left(\frac{p}{c}\right). \tag{14}$$

再将(13) 式代入(14) 式得到

$$S(\nu) = \frac{2}{c}\int_{-\infty}^{\infty} [I(p) - A]\exp\left(\frac{\mathrm{i}2\pi p\nu}{c}\right)\mathrm{d}p. \tag{15}$$

最后,我们看到光波的频率分布 $S(\nu)$ 可以由 $[I(p) - A]$ 的傅里叶逆变换得到,而 $[I(p) - A]$ 可以用移动反射镜 M_2 的过程中多次抽样测量的数据来充分逼近. 与前面的假设相对应,我们设 $S(-\nu) = S(\nu)$,$S(-\nu)$ 没有物理意义,我们将它略去.

在上一节中,我们介绍了相干长度的物理意义,当迈克耳孙两臂的长度差超出相干长度时,干涉现象消失. 事实上,相干长度取

决于光波中所包含的频率成分.如果光波中只包含单一频率,我们将得到无限长的正弦波序列,时空中任意一点的相位都是确定的,相干长度为无穷.反之,如果光波中包含较为丰富的频率成分,即频带$\Delta\nu$较宽,对应的光波就是"无序"的,亦即相干长度很短.根据傅里叶变换的一般规律,时间t的"不确定度"τ和频率的"不确定度"$\Delta\nu$满足测不准关系

$$\Delta\nu \cdot \tau \sim 1. \tag{16}$$

由于相干长度$l = \tau \cdot c$,所以有

$$\Delta\nu \cdot l \sim c. \tag{17}$$

(16)、(17)式深刻地反映了相干长度l或相干时间τ与带宽$\Delta\nu$的相互制约的关系,当$S(\nu)$具有高斯分布时,

$$\Delta\nu \cdot \tau = 1. \tag{18}$$

关于测不准关系的证明请参见节1.2,或文献[6].

3.7 投影显示的消像素技术

当使用液晶显示器LCD(参见第七章)进行投影成像时,LCD上的像素结构就会出现在投影屏上.消除投影像中的像素结构的工作请参见文献[7]和[8],LCD是矩阵寻址的空间光调制器,它的像素结构是必不可缺的.我们的研究目的则是设法消除投影像中的像素结构.众所周知,在$4f$系统或其他成像系统频率平面中心的小孔可以通过相干光图像处理消除栅网结构(参见图2.8).由于像素的周期结构,在频率平面上出现一系列傅里叶谱项.这正是原始图形与周期结构函数的乘积经傅里叶变换后,在频率平面上形成的图形的谱与δ函数列阵的卷积(参见节1.1).每一个谱项都只是中心谱项在不同位置的"复现",因此只要在频率平面上放置小孔滤波器,仅让一个谱项(例如零级谱项)通过,就可以消除像素结构,见图3.6.然而在这一过程中大部分能量都被滤波器拦去,输出像十分暗淡.

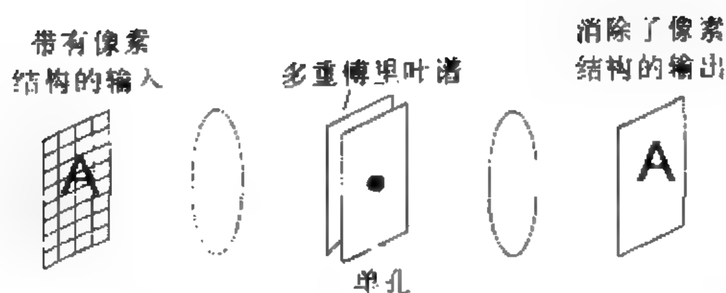


图 3.6　利用小孔滤波消除像素结构

下面的现象是很有趣的：选取任意一个谱项通过小孔，产生的图像均位于同一位置. 如果两个谱项通过两个小孔，产生的图像上就可以看到杨氏条纹. 如果所有的谱项一起通过滤波(事实上不放任何滤波器)，产生的干涉条纹就综合形成了像素的结构. 上面讲过，如果只让一个谱项通过滤波小孔，尽管像素结构消除了，但能量损失太大. 为了弥补这一缺点，可以让所有的谱项都通过傅里叶变换平面. 既然像素结构是由干涉效应引起的，我们可以让不同的谱项通过不同厚度的透明的相位片，以获得不同的延迟，参见图 3.7. 只要相位的延迟大于相干长度，各谱项间的相干性就被破坏，结果像素结构就消除了，而强度并不受到影响. 对于白光光源，由于频谱很宽，典型的相干长度仅 10 μm 左右.

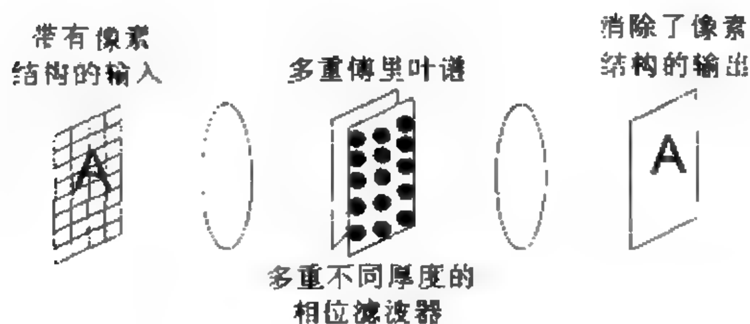
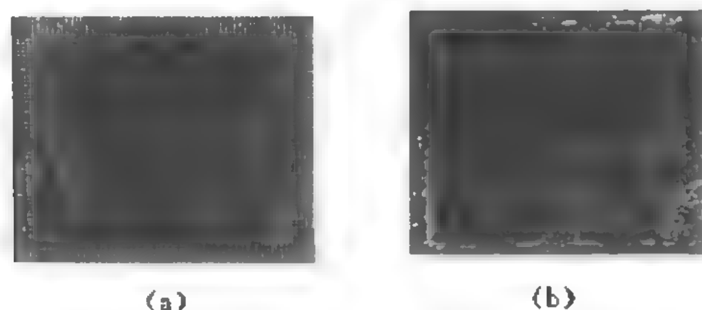


图 3.7　利用不同相位延迟的相位滤波消除像素结构

上述技术显著地改善了液晶投影显示的成像质量. 特别是一些分辨率较低的投影仪, 利用这一技术改善了像质, 使它的投影像看起来好像是高分辨率的投影仪. 图 3.8 是消像素结构的实验结果. 图(a) 中带有像素结构的投影像, 频率平面上没有相位滤波片. 图(b) 表示在频率平面上加相位滤波器后, 消除了像素结构.



(a) (b)

图 3.8

3.8 计算层析技术

运用 X 射线的计算层析技术, 又称 CT 断层扫描, 向来认为不属于光学信息处理的范畴. 然而我们认为有必要对它进行综合评述, 因为这一技术所采用的正是光学信息处理常用的傅里叶变换. 如果认为 X 射线是"超短光波", 那么层析技术也可认为是光学信息处理的应用.

X 射线 CT 用于获取人体的剖面图像. 注意通常的 X 射线的图像仅仅是投影, 而不是图像本身.

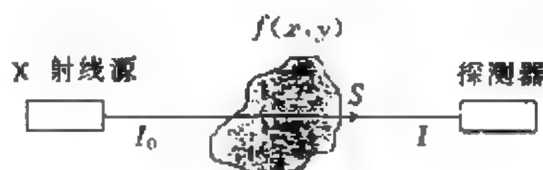


图 3.9 X 射线通过一个物体

参见图 3.9, 设 X 射线穿透一个物体, 透过率或光衰减率用 $f(x,y)$ 表示. 探测到的 X 射线的强度分布为

$$I = I_0 \exp\left\{-\int f(x,y)\mathrm{d}S\right\}, \tag{1}$$

式中 I_0 为初始光强.

(1) 式可改写为

$$\int f(x,y)\mathrm{d}S = -\ln \frac{I}{I_0}, \tag{2}$$

尽管可以测得积分$\int f(x,y)\mathrm{d}S$,但我们并不能得到被积函数 $f(x,y)$. 然而真正需要的还是 $f(x,y)$,即断层图像. 层析术的基本思路,是从 $f(x,y)$ 的傅里叶变换 $F(u,v)$ 入手. $F(u,v)$ 可表为

$$F(u,v) = \iint_{-\infty}^{\infty} f(x,y)\exp[-\mathrm{i}2\pi(ux+vy)]\mathrm{d}x\mathrm{d}y. \tag{3}$$

设 $v=0$,上式化作

$$\begin{aligned} F(u,0) &= \iint_{-\infty}^{\infty} f(x,y)\exp(-\mathrm{i}2\pi ux)\mathrm{d}x\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x,y)\mathrm{d}y\right]\exp(-\mathrm{i}2\pi ux)\mathrm{d}x. \end{aligned} \tag{4}$$

上式可进一步改写为

$$F(u,0) = \int_{-\infty}^{\infty} p(x)\exp(-\mathrm{i}2\pi ux)\mathrm{d}x, \tag{5}$$

式中

$$p(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y. \tag{6}$$

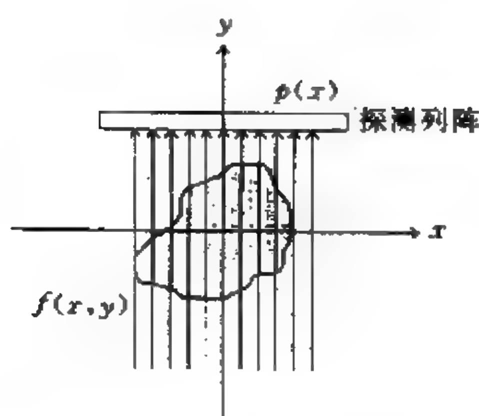


图 3.10 X 射线平行于 y 轴穿透物体,产生函数 $p(x)$

参见图 3.9 及(2) 式,(6) 式所示的 $p(x)$ 可以用 X 射线沿平行于 y 的方向的投影得到,见图 3.10.

如果 $p(x)$ 用图 3.10 的方法得到,$F(u,0)$ 就可以根据公式(5) 计算出来,$F(u,0)$ 是沿 u 轴方向的傅里叶变换. 为了获得更多的信息,可以把坐标轴 (u,v) 转动一个小角度 α 得到 (u',v'),相应地把 (x,y) 旋转 α 角得到

(x',y')，而物体不转动，见图 3.11. 于是我们得到 $p(x')$，并采用同样的手续获得 $F(u',v'=0)$；再把坐标系转过 α，我们又获取 $F(u'',v''=0)$，总共旋转 180°，就得到傅里叶频域中的一系列数据，参见图 3.12. 图中旋转坐标系测出 $p(x)$，$p(x')$，$p(x'')$，…，由此计算出 $F(u,v)$，物体保持不动.

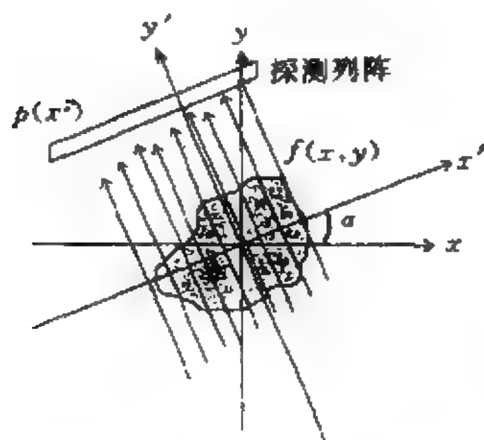


图 3.11 用倾斜的 X 射线穿透物体生成 $p(x')$

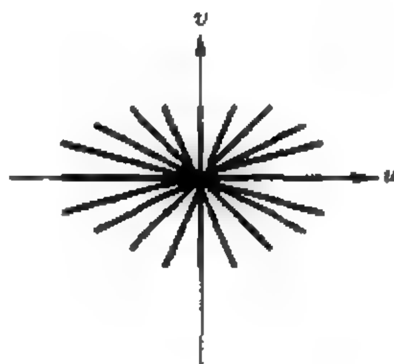


图 3.12

根据图 3.12 所示的方案得到 $F(u,v)$ 的数据集合后，就可以通过傅里叶逆变换得到 $f(x,y)$. 有两种不同的处理方法：第一种方法通过内插，得到在直角坐标系中的 $F(u,v)$，然后进行傅里叶逆变换；第二种方法直接在极坐标下进行傅里叶变换，因而免去了在傅里叶域中的内插手续. 直角坐标系下的逆变换表达式为

$$f(x,y)=\iint_{-\infty}^{\infty} f(u,v)\exp[\mathrm{i}2\pi(ux+vy)]\mathrm{d}u\mathrm{d}v. \tag{7}$$

该式可以改写为极坐标中的表达式

$$f(x,y)=\int_0^{2\pi}\int_0^{\infty} F(\rho\cos\phi,\rho\sin\phi)\exp[\mathrm{i}2\pi\rho(x\cos\phi+y\sin\phi)]\rho\mathrm{d}\rho\mathrm{d}\varphi, \tag{8}$$

其中

$$u=\rho\cos\phi, \tag{9}$$

$$v = \rho \sin \phi. \tag{10}$$

无须内插，就可以从不同旋转角度下 $F(u,v)$ 的数据，运用(8)式得到 $f(x,y)$. 公式(5)、(7)、(8)均可用计算机算出，其中的积分用求和代替. 图 3.13 是完整的CT图，脊椎和肋骨呈现出明亮的白色，图像左边的大块的组织是肝脏，右上角可看到胃里的液体和气体.



图 3.13　人体的 CT 图像[9]

3.9　结　　论

在本章中，我们对比了非相干光与相干光的成像过程与空间滤波. 非相干光学传递函数(OTF) 为相干光学传递函数(CTF) 的自相关，非相干系统的本性就是低频滤波. 我们简要分析了光的相干性，并把傅里叶分析方法运用到傅里叶变换光谱仪及计算层析术(CT) 中去. 在本章中我们还介绍了消除液晶显示技术中像素结构的方法.

参 考 文 献

[1]　A. S. Marrathay, Elements of Optical Coherence Theory, John Wiley &Sons, New York, 1982.

[2]　羊国光，宋菲君，高等物理光学，中国科技大学出版社，1991.

[3]　J. W. Goodman, Introduction to Fourier Optics, McGraw Hill, San Franscisco,

1968.

[4] A. W. Lohmann and H. W. Werlich, Incoherent matched filtering with Fourier holograms, *Appl. Opt.*, **10**(1971), 670-672.

[5] A. Vander Lugt, Signal detection by complex spatial filtering, *IEEE Trans. Inf. Theo.*, IT~**10**(1964), 139-145.

[6] 曾谨言,量子力学,科学出版社,1984.

[7] S. Jutamulia, S. Toyoda, Y. Ichibashi, Removal of pixel structure in liquid crystal projection display, *Proc. SPIE*, **2407**(1995), 168-176.

[8] X. Yang, S. Jutamulia, N. Li, Liquid-crystal projection image depixelzation by spatial phase scrambling, *Appl. Opt.*, **35**(1996), 4577-4580.

[9] K. Iizuka, Engineering Optics, Springer-Verlag, Berlin, 1985, 295.

第四章 光学图像识别

4.1 图像识别和光学相关器

很久以来，人们一直在研究能够识别物体的机器，这种机器能代替人们从事枯燥乏味的重复性劳动及危险性的工作. 例如，字符识别机能代替邮递员分拣邮件；自动签名或指纹识别机能代替工作人员检验签字或指纹. 识别机比人更加可靠，因为它不会疲劳，而工作人员却会因为长期工作的疲劳而造成错判或漏检. 在军事上，首先用图像识别系统辨认对方的目标，诸如导弹、车辆、建筑，而后用武器自动摧毁经过识别的目标. 在医学上，图像识别技术则用于识别某一类特定的细胞，然后进行计数.

自动图像识别的基本结构是光学相关器[1]. 如第一章所述，两个实函数 $f(x,y)$ 和 $g(x,y)$ 的相关在数学上定义为

$$c(\alpha,\beta) = \iint_{-\infty}^{\infty} f(x,y)g(x-\alpha,y-\beta)\mathrm{d}x\mathrm{d}y. \tag{1}$$

为了运算方便，我们改变了宗量的符号. 这一变化不会影响相关运算的物理意义. 当 $\alpha=0,\beta=0$ 时相关函数变成内积：

$$c(0,0) = \iint_{-\infty}^{\infty} f(x,y)g(x,y)\mathrm{d}x\mathrm{d}y. \tag{2}$$

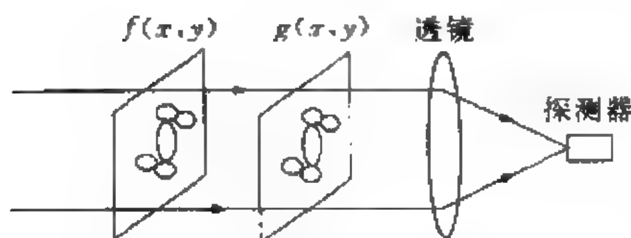


图 4.1 光学图像识别器图示

如果 $f(x,y)$ 和 $g(x,y)$ 是两个光学模板或透明片，(2) 式表示两个透明物体是互相交叠的，光束依次通过两个物体，并用透镜聚焦到探测器

上进行积分，如图 4.1 所示[2]. 显然，仅当 $f(x,y)$ 和 $g(x,y)$ 全同时，光强的积分 $c(0,0)$ 才达到极大值.

这样一来，只要测出光强的积分，就可以判定图形 $f(x,y)$ 和 $g(x,y)$ 的相似性. 请注意，这一识别过程把复杂的二维图形化简为一个点，如果 $g(x,y)$ 是参考信号，则输入信号 $f(x,y)$ 与参考信号的相似性的度量转化为一个点的强度测量.

由此可见，相关器可以用作图形识别器. 相关器的局限性也是十分明显的，它仅能判别两个完全相同的图形，倘若一个图形相对于另一个图形转过一个角度，或二者的比例有所不同，即便两个图形完全相同，用相关器是无法识别的. 其结果，相关器很容易识别印刷体字符，却无法识别手写体字符. 此外，它容易识别人脸的照片，对真人的脸部的识别却无能为力. 总而言之，相关器可用来识别具有明确定义的物体. 定义不明确的物体则应借助于神经网络来识别，可参见第八章.

4.2　非相干识别器

在图 4.1 所示的系统中，$f(x,y)$ 和 $g(x,y)$ 必须准确地叠合，以产生最大的积分强度峰. 一旦两个图形有所错位，它们就不再精确地重合，从而探测到的强度峰迅速消失，$f(x,y)$ 就不能被识别. 由上节公式(1)，参考信号可以沿 x 或 y 方向平移[2]，例如可以通过计算机控制的步进电机实施平移，从而相关峰的积分强度就是 $g(x,y)$ 的位置 (α,β) 的函数，测得的相关函数 $c(\alpha,\beta)$ 将指示输入信号 $f(x,y)$ 与参考信号 $g(x,y)$ 是否相同，并给出它们精确重合时的位置.

另一光学系统见图 4.2[2]，它不用准直光束，而用一个扩展的面光源以同时产生许多个准直光束. 面光源可以看成是大量均匀分布的点光源的集合，如果把面光源放在透镜的焦平面上，每个点源将生成一束平行光. $f(x,y)$ 和 $g(x,y)$ 分开一个距离 d，它们将

分别收集通过透明片的平行光束，而其他部分的光束则被阻挡. 如果两个图形透明部分相对于这一平行光束是重合的，它通过第二个透镜后就在焦平面上产生一个亮斑，它正是两个图形的相关峰，并指示其位置. 倘若输入平面上存在不止一个图形，则有不止一个平行光束通过这些图形，在输出平面上出现多个相关斑.

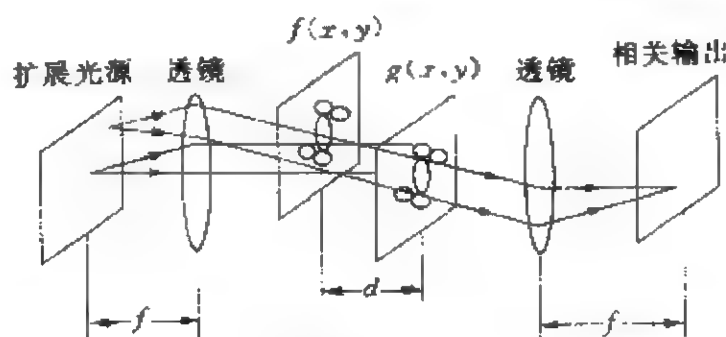


图 4.2 使用扩展光源的非相干光学相关器

相关输出 $c(\xi,\eta)$ 则可表为

$$c(\xi,\eta)=\iint_{-\infty}^{\infty} f(x,y)g\left(x-\frac{\xi d}{f},y-\frac{\eta d}{f}\right)\mathrm{d}x\mathrm{d}y.$$

相关输出平面为第二个透镜的后焦面.

很明显，图 4.2 所示的非相干光学相关器有一个缺点，即输出平面上的相关峰强度很弱，因为相关峰是扩展光源上的一小部分所产生的平行光束形成的.

4.3 Vander Lugt 相关器

Vander Lugt 相关器(VLC)[3] 应用了透镜的傅里叶变换性质(参见第二章)，为了施行节 4.1(1) 式所示的相关运算，VLC 对于 $F(u,v)G^*(u,v)$ 进行傅里叶逆变换，这里 F 和 G 分别是 f 和 g 的傅里叶变换. VLC 的基本概念可以由图 4.3 及图 4.4 表征.

首先必须把 $g(x,y)$ 的傅里叶复共轭谱 $G^*(u,v)$ 存储起来. $g(x,y)$ 是实函数，但 $G^*(u,v)$ 通常是复函数，所以必须存储在全

息图中，参见图 4.3. 将参考图形 $g(x,y)$ 置于透镜的输入平面上，在透镜的后焦面即谱平面上出现它的谱 $G(u,v)$，再添加一个离轴的参考光 $R(u,v)$ 共同照射谱平面，设在谱面上放置一个感光胶片，并假定胶片的显影、定影过程是正的线性过程，则最后得到的全息滤波片为

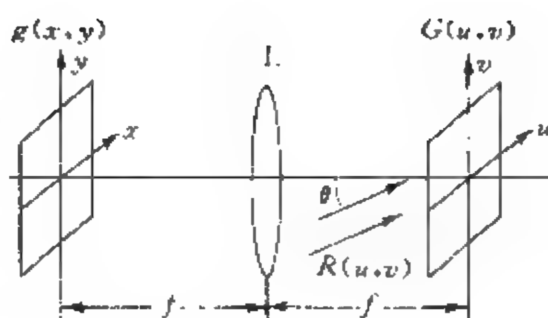


图 4.3 VLC 复数滤波器的记录

$$T(u,v) = |G(u,v) + R(u,v)|^2, \tag{1}$$

式中 $R(u,v)$ 是一个单位光强的倾斜入射的平行光，数学上可表为

$$R(u,v) = \exp[iau] \tag{2}$$

并有

$$a = \frac{2\pi\sin\theta}{\lambda}, \tag{3}$$

θ 是参考光 R 与光轴的夹角，λ 为波长. 将公式(2) 代入(1)，得到

$$T(u,v) = |G(u,v)|^2 + 1 + G(u,v)\exp[-iau] + G^*(u,v)\exp[iau], \tag{4}$$

识别过程由图 4.4 所示.

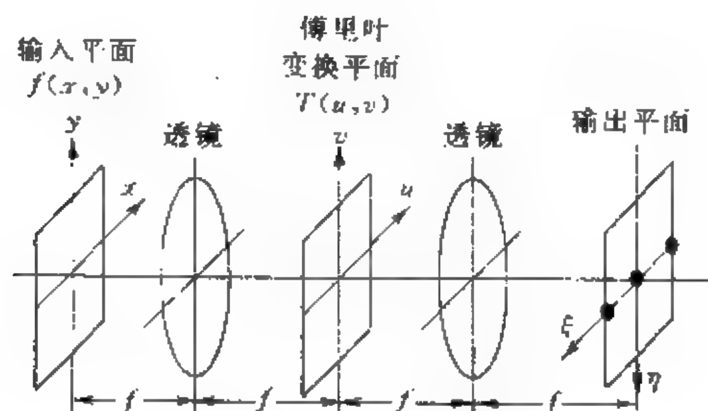


图 4.4 VLC 的识别过程

(4) 式所示的透过率函数 $T(u,v)$ 称为复数匹配滤波器,将它置于图 4.4 所示 $4f$ 系统的傅里叶变换平面上. 再将函数 $f(x,y)$ 置于输入平面上,在傅里叶平面上将产生 f 的谱 $F(u,v)$,经过 $T(u,v)$ 滤波作用,紧贴 $T(u,v)$ 后面的场为

$$F(u,v)T(u,v) = (|G(u,v)|^2 + 1)F(u,v) + F(u,v)G(u,v)\exp[-\mathrm{i}au] + F(u,v)G^*(u,v)\exp[\mathrm{i}au], \tag{5}$$

第二个透镜则在输出平面(ξ,η) 上形成 $F(u,v)T(u,v)$ 的傅里叶逆变换. 由第二章我们知道上式第一项位于输出平面中心,形成 0 级谱,而 $F(u,v)G(u,v)$ 及 $F(u,v)G^*(u,v)$ 的逆变换将分别出现在输出平面上$(b,0)$ 及$(-b,0)$ 处,其中

$$b = f\sin\theta. \tag{6}$$

为了更清楚地阐明上面所讲的问题,我们重新推导 $F(u,v)G^*(u,v)\exp[\mathrm{i}au]$ 的傅里叶逆变换:

$$c(\xi,\eta) = \iint_{-\infty}^{\infty} F(u,v)G^*(u,v)\exp[\mathrm{i}au]\exp\left[\mathrm{i}\,\frac{2\pi}{\lambda f}(\xi u + \eta v)\right]\mathrm{d}u\mathrm{d}v, \tag{7}$$

其中 $F(u,v)$ 及 $G^*(u,v)$ 可以表为

$$F(u,v) = \iint_{-\infty}^{\infty} f(\alpha,\beta)\exp\left[-\mathrm{i}\,\frac{2\pi}{\lambda f}(\alpha u + \beta v)\right]\mathrm{d}\alpha\mathrm{d}\beta, \tag{8}$$

$$G^*(u,v) = \iint_{-\infty}^{\infty} g^*(-p,-q)\exp\left[-\mathrm{i}\,\frac{2\pi}{\lambda f}(pu + qv)\right]\mathrm{d}p\mathrm{d}q, \tag{9}$$

代入(7) 式得到

$$c(\xi,\eta) = \iint_{-\infty}^{\infty} f(\alpha,\beta) \iint_{-\infty}^{\infty} g^*(-p,-q) \iint_{-\infty}^{\infty} \exp\left\{\mathrm{i}\,\frac{2\pi}{\lambda f}[(\xi+b-\alpha-p)u+(\eta-\beta-q)v]\right\}\mathrm{d}u\mathrm{d}v\mathrm{d}p\mathrm{d}q\mathrm{d}\alpha\mathrm{d}\beta, \tag{10}$$

其中

$$b = fa\lambda/2\pi, \tag{11}$$

将(3) 式代入得到

$$b = f\sin\theta. \tag{12}$$

这样一来,(10) 式进一步化作

$$\begin{aligned} c(\xi,\eta) = & \iint_{-\infty}^{\infty} f(\alpha,\beta) \iint_{-\infty}^{\infty} g^*(-p,-q)\delta(\xi + b \\ & - \alpha - p, \eta - \beta - q)\mathrm{d}p\mathrm{d}q\mathrm{d}\alpha\mathrm{d}\beta \\ = & \iint_{-\infty}^{\infty} f(\alpha,\beta) g^*(\alpha - (\xi + b), \beta - \eta)\mathrm{d}\alpha\mathrm{d}\beta. \end{aligned} \tag{13}$$

如果 $g(x,y)$ 是实的,则上式化为

$$c(\xi,\eta) = \iint_{-\infty}^{\infty} f(\alpha,\beta) g(\alpha - (\xi + b), \beta - \eta)\mathrm{d}\alpha\mathrm{d}\beta, \tag{14}$$

它表示 f 和 g 的相关,在 ξ 方向平移 $-b$. 注意参考光波 $R(u,v) = \exp[\mathrm{i}au]$ 的傅里叶逆变换本身就是位于 $(-b,0)$ 的 δ 函数,因而相关输出呈现在参考光的焦点处.

从物理上看,全息图由 $g(x,y)$ 的傅里叶变换谱 $G(u,v)$(物光) 和参考光 $R(u,v)$ 形成. 而物光和参考光本来是相对而言的,当第二次输入的信号 $f(x,y)$ 与 $g(x,y)$ 一致时,它的谱 $F(u,v)$ 和 $G(u,v)$ 也应一致,它照射全息图,必然重建了参考光 $R(u,v)$,从而在输出平面上出现 $R(u,v)$ 的焦点,这是一个由物光重建参考光的过程. 显然,仅当 f 准确地与 g 相同时才能重建参考光,所以相关峰是 f 和 g 一致性的标志.

4.4 实时 Vander Lugt 相关器

VLC 的基本装置如图 4.5 所示. 图中 L_1 为显微物镜;L_2 为准直透镜;L_3,L_4 为傅里叶变换透镜;PH 为针孔;BS 为分光镜;M 为

反光镜；MSF 为匹配滤波器；AD 为探测列阵；OB 为物光；RB 为参考光.

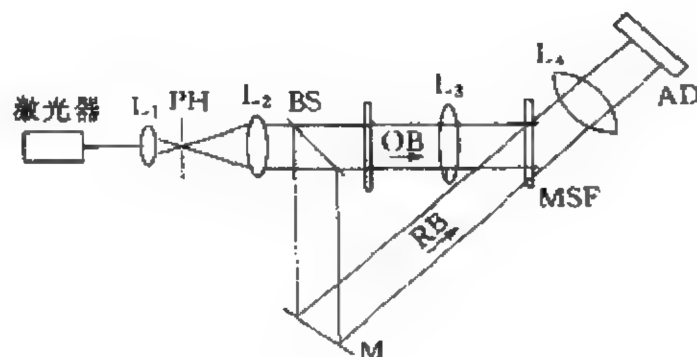


图 4.5 VLC 的基本设备

激光束首先由显微物镜 L_1 聚焦，通过针孔滤波后，再由准直镜 L_2 形成平行光束照射输入平面，由于针孔滤波，光束的截面强度是均匀的. 平行光由分光镜 BS 分成物光 OB 及参考光 RB. 当在输入平面上放置参考信号 $g(x,y)$ 时，我们就可以用全息材料记录匹配滤波器 MSF，输入平面和 MSF 平面分别是 L_3 的前、后焦面. 处理后的全息图重新放在 MSF 处.

当 MSF 精确地复位后，挡去参考光 RB，在输入平面上放置信号 $f(x,y)$. 由于相关信号（即由物光重建的参考光）是准确地沿原来的参考光 RB 方向传播的，所以我们把另一个傅里叶变换透镜 L_4 放在 RB 中，与 RB 的光轴垂直. 最后用探测列阵 AD（例如 CCD）来探测相关峰.

在上面的装置中，假定 $f(x,y)$ 和 $g(x,y)$ 是由透明片来显示的. 在实时图形识别中，我们既不能使用胶片来输入图形，更不可能对胶片进行冲洗手续. 为了实现实时图像识别，我们用空间光调制器 SLM 来输入图形. 关于空间光调制器的详细介绍，请参见第七章.

图 4.5 是一个使用透射型 SLM（例如液晶显示器 LCD）的实时图像显示系统. 图 4.6 是使用反射型 SLM（例如液晶光阀

LCLV，均参见第七章）的相应光路．在前一情形下用 CCD 实时拍摄物体，并通过图像处理器把被拍摄的物体像显示在 LCD 上．图 4.6 与图 4.5 对应的系统略有区别．

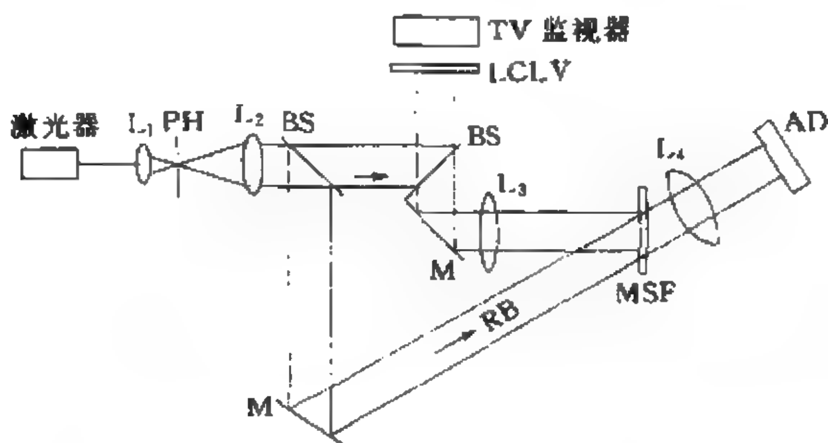


图 4.6　使用 LCLV 及电视显示终端实现实时 VLC 的基本光学系统

在图 4.6 中，由摄像机拍摄的图像由 TV 终端显示，该图像直接照射 LCLV，或用透镜投影到 LCLV 的写入端面，激光束则在 LCLV 的读出端面上反射，并读出信号，即 VLC 的输入信号 $f(x,y)$．

图 4.7 是一个实时图像识别的例子．工厂中移动的物体由实

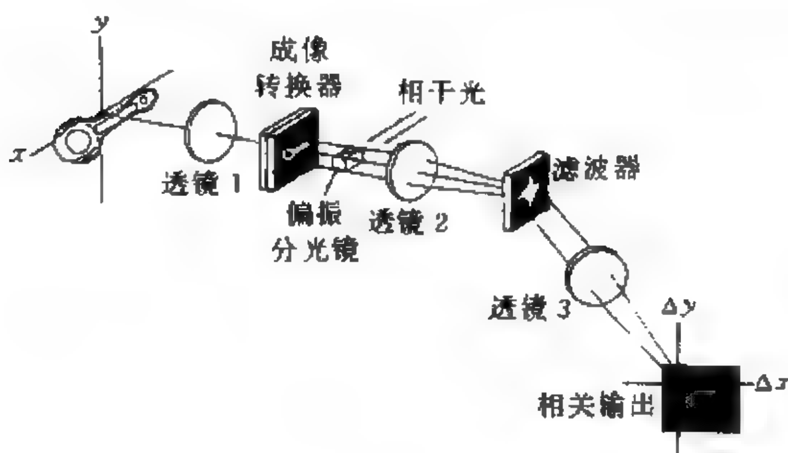


图 4.7　应用 LCLV 进行实时相关识别[4]

时相关器识别并跟踪[4]，在传送带上的物体由功率为 200 W 的钨丝灯照亮，并通过传送带上方的反射镜系统，投影到LCLV的写入端，该系统既无摄像机，又不用终端显示器. LCLV 的读出端则作为相关器的输入. 图 4.7 左面的物体出现时，在相关平面上就出现亮斑，亮斑的位置与物体的位置具有线性的关系.

4.5 Vander Lugt 相关器的小型化

Vander Lugt 相关器已成功地实现了小型化，可以装在飞机上，甚至装在导弹的头部. 文献[5] 报道了这类小型化的 VLC. 为了装在小盒子内，光路几次折叠、多处交叉. 底板尺寸为 31 cm × 23 cm，光束直径 2.5 cm，外壳高度 15 cm，内装驱动电路(见图 4.8).

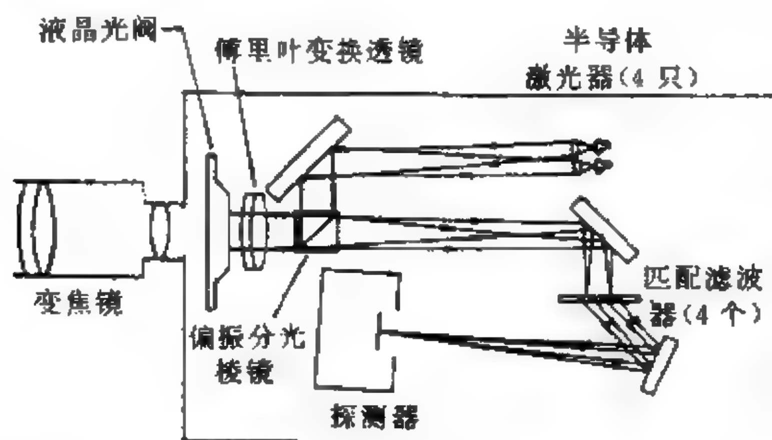


图 4.8　小型化实时相干光学相关器光学系统图[5]

小型化 VLC 用半导体激光器作光源. 四只波长为 830 nm、峰值功率为 10 mW 的半导体激光器辐射四束独立的光束，在变换平面上生成四个独立寻址的匹配滤波器 MSF. 激光束的发散角为 20° × 40°，用微型透镜部分准直. 系统中不用针孔滤波器和显微物镜.

待识别的物体通过变焦物镜成像到 LCLV 的写入端面. 系统其余部分是一个典型的 VLC，只是光路用反射镜加以折转以节约

空间. 四个激光器辐射的光束经折转后，由偏振分光棱镜反射，反射光为 s 偏振，它就是 LCLV 的读出光，读出的图像信号为 p 偏振分量，透过偏振分光棱镜进入识别系统. 在棱镜和 LCLV 之间有傅里叶变换透镜，在透镜的后焦面上放置匹配滤波器 MSF. 由于 MSF 是用会聚参考光制作的，这样就可以省去一个逆变换透镜. 一旦输入信号与参考信号相同，系统立即产生相关输出，由 CCD 接收.

ZOOM 透镜可以调节焦距，使目标的大小与参考信号匹配. LCLV 的响应时间为 0.1 s，它限制了识别速度的进一步提高.

匹配滤波器是用另一个系统制作的，制作时使用 He-Ne 激光 (633 nm)，因为全息记录材料对 830 nm 的红外光不敏感. 这样一来，记录过程和相关识别过程的波长就有 30% 的差别. 这一差别产生三个效应：傅里叶变换的尺度放大，相关信号的衍射角增大，并在相关平面上引入了附加的像差. 前两个效应可以通过几何光学的放大率控制来消除. 图 4.9(a) 为一个坦克模型，(b) 表示对应的匹配滤波器 MSF，(c) 乃是相关信号. 据报道，小型 VLC 在实验室和现场的性能都令人满意.

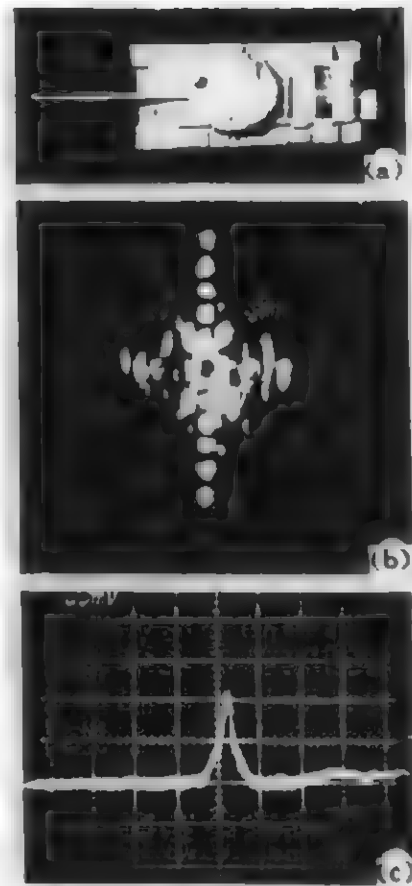


图 4.9　坦克模型的相关识别实验[5]
(a) 输入目标
(b) 输入目标的 MSF
(c) 相关信号

图 4.10 是近年报道的小型化 VLC 示意图，用三个普罗 (Porro) 棱镜和一个分光棱镜胶合成固化的棱镜组，整个光学系统高 45 mm，对角线长 88 mm，内部性能与上述系统类似，但只有

…个通道. 该系统已装备在导弹头部. 图 4.11 是该系统拆开后的实物图.

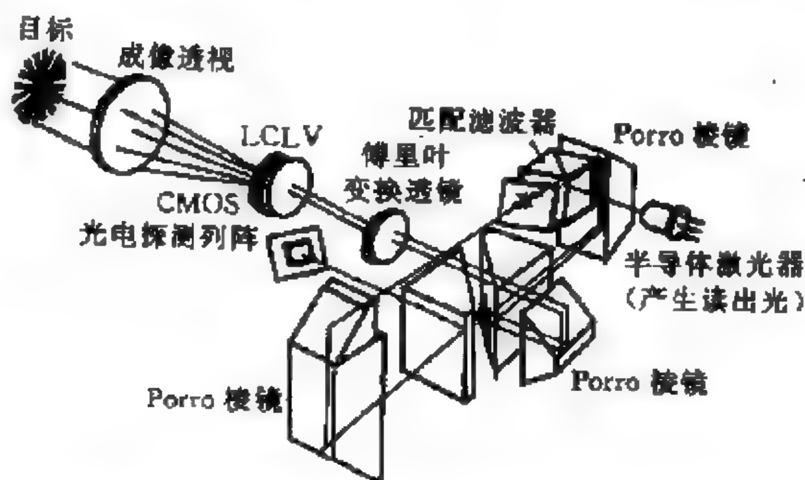


图 4.10 固态光学相关器部件分解图[6]

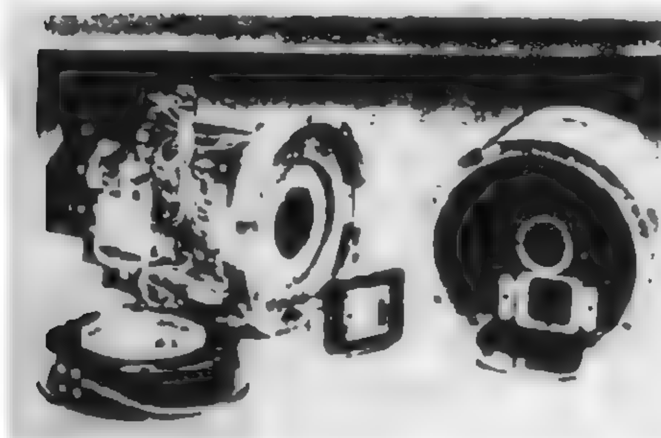


图 4.11 部分拆开的相关器部件图[8]

指纹是最先用光学相关器加以识别的图形. 实时指纹识别仪已有报道[7]. 指纹可用作身份的特征,因为事实上任何两个人的指纹都不相同. 每个人都可以用自己的指纹制作 MSF,并存储在身份证件中,例如各种信用卡、钥匙等. 余下的问题只是如何用简单而实时的方法把待识别的指纹图案引入识别器,与证件中的 MSF 加以对比识别. 图 4.12 为反射型实时指纹识别器[7],指纹借助于全反射效应导入系统. 系统上方有一个直角棱镜,半导体激光器的

光束在棱镜斜面发生全内反射进入系统. 当用手指压在斜面上时,指纹图案破坏了全内反射条件,调制了激光束,形成输入图像. 匹配滤波器也是反射形的,图中没有表示出来.

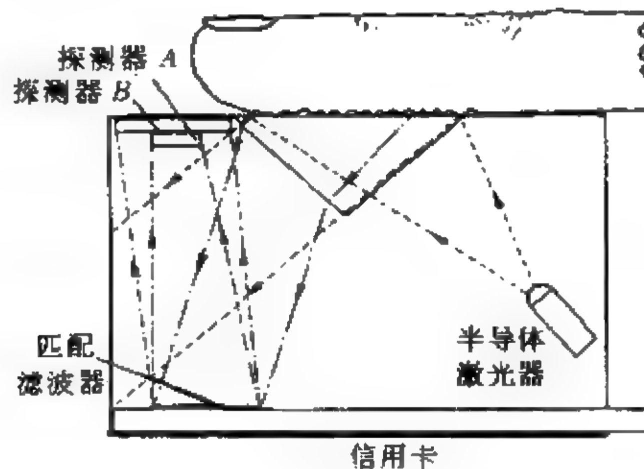


图 4.12 反射型 MSF 实时指纹识别器[7]

4.6 旋转不变 Vander Lugt 相关器

节 4.1(1) 式表明,仅当输入图像与参考图像的形状、比例完全相同时才有相关输出,即使两个图像形状一致,但比例不同(两图形相似)或相对旋转一个角度,相关输出都会急剧下降,无法靠相关峰来识别. 多年来,人们一直在研究旋转和缩放不变的 Vander Lugt 相关器. 我们将介绍一个最早提出来的利用圆谐函数展开实现旋转不变 VLC 的方案[8].

输入函数 $f(x,y)$ 可以用极坐标表为 $f(r,\theta)$,并用指数函数展开:

$$f(r,\theta)=\sum_{M=-\infty}^{\infty}f_M(r)\exp[\mathrm{i}M\theta], \tag{1}$$

$$f_M(r)=\frac{1}{2\pi}\int_0^{2\pi}f(r,\theta)\exp[-\mathrm{i}M\theta]\mathrm{d}\theta. \tag{2}$$

这样一来,旋转 ϕ 角的目标函数可表为

$$f(r,\theta+\phi)=\sum_{M=-\infty}^{\infty}f_M(r)\exp[\mathrm{i}M\phi]\exp[\mathrm{i}M\theta]. \qquad (3)$$

在原点(0,0)的相关在极坐标中的表达式为

$$c(\phi)=c_\phi(0,0)=\int_0^{\infty}r\mathrm{d}r\int_0^{2\pi}f(r,\theta+\phi)f^*(r,\theta)\mathrm{d}\theta. \qquad (4)$$

以(1)式代入(4)式，得到

$$\begin{aligned} c(\phi) &= \sum_{M=-\infty}^{\infty}\sum_{M'=-\infty}^{\infty}\Big\{\exp[\mathrm{i}M\phi]\int_0^{\infty}f_M(r)f_{M'}(r)r\mathrm{d}r \\ &\quad\times\int_0^{2\pi}\exp[\mathrm{i}(M-M')\theta]\mathrm{d}\theta\Big\} \\ &= 2\pi\sum_{M=-\infty}^{\infty}\exp[\mathrm{i}M\phi]\int_0^{\infty}|f_M(r)|^2r\mathrm{d}r. \end{aligned} \qquad (5)$$

在推导中我们运用了极坐标中的 δ 函数的表达式

$$\frac{1}{2\pi}\int_0^{2\pi}\exp[\mathrm{i}(M-M')\theta]\mathrm{d}\theta=\delta_{MM'}. \qquad (6)$$

(5)式所示的相关函数包含了各级圆谐函数分量的贡献，当旋转角 ϕ 变化时，$c(\phi)$ 显然不满足旋转不变的条件. 这样我们得到结论：当参考信号中包含多个(大于一个)圆谐函数分量时相关输出是旋转可变的.

显然，如果参考信号中仅包含一个(某一级)圆谐函数分量时就可以实现旋转不变. 设

$$f_R(r,\theta)=f_M(r)\exp[\mathrm{i}M\theta], \qquad (7)$$

代入(5)式，得到相关输出

$$c_M(\phi)=2\pi\exp[\mathrm{i}M\phi]\int_0^{\infty}r|f_M(r)|^2\mathrm{d}r, \qquad (8)$$

相关函数的强度

$$|c_M(\phi)^2|=\left|2\pi\int_0^{\infty}r|f_M(r)|^2\mathrm{d}r\right|^2. \qquad (9)$$

它与目标图形的旋转角 ϕ 无关，因而是旋转不变的.

这一旋转不变的相关识别过程，显然强烈地依赖于极坐标系原点的选择及圆谐函数分量的级 M 的选择. 一般的原则是把原点

选在图形的对称中心或大致的中心附近，并选择较低级次的圆谐函数分量作为参考信号.

在图4.13(a)中，字母E作为目标函数，并用"×"指出圆谐函数的展开中心；图(b)给出 $M=1$ 时的圆谐函数分量的振幅；图(c)和(d)分别是(b)中圆谐函数的实部和虚部. 与该复圆谐函数匹配的空间滤波器 MSF 用计算机产生的全息图(CGH, computer generated holographic)制作. 图 4.14 给出旋转不变的相关峰，指示了三个不同取向的字母E的存在，在识别过程中取了适当的阈值.

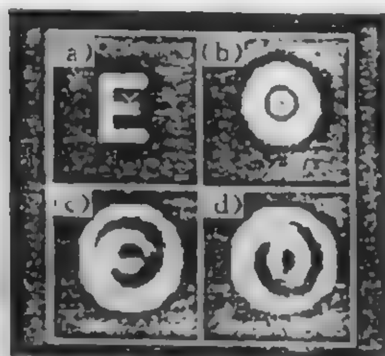


图 4.13　第一级圆谐函数分量[8]



图 4.14　对字母 E 识别的旋转不变相关输出

4.7　比例不变 Vander Lugt 相关器

比例不变又称尺度缩放不变，意思是说目标图形与参考图形相似时，仍有相关峰输出. 引入下面的变量代换：

$$\left.\begin{aligned}\xi &= \ln[x],\\ \eta &= \ln[y].\end{aligned}\right\} \tag{1}$$

图形函数借助于宗量 ξ,η 表为

$$f(\xi,\eta)=f(\ln[x],\ln[y]). \tag{2}$$

对原函数进行 a 倍的比例缩放的结果为

$$f_a(\xi,\eta)=f(\ln[ax],\ln[ay]). \tag{3}$$

上式可进一步写为

$$f_a(\xi,\eta)=f(\xi+\xi_1,\eta+\eta_1), \tag{4}$$

式中

$$\xi_1=\ln[a], \qquad \eta_1=\ln[a]. \tag{5}$$

注意，$\xi_1=\eta_1$.

如果我们对变换后的图形函数 $f(\xi,\eta)$ 与 $f(\xi+\xi_1,\eta+\eta_1)$ 进行相关识别运算，显然有相关峰输出. 而 $f(\xi,\eta)$ 与 $f(x,y)$ 对应，$f(\xi+\xi_1,\eta+\eta_1)$ 与 $f(ax,ay)$ 对应，恰恰是原来的图形函数与比例缩放后的图形函数的相关，相关峰的位移 $(-\xi_1,-\eta_1)$ 指示了缩放的系数

$$a=\exp[\xi_1]=\exp[\eta_1]. \tag{6}$$

在上述处理过程中，我们不能以 $f(ax,ay)$ 作为输入信号，而必须用 $f(\xi+\xi_1,\eta+\eta_1)$ 作为输入信号，输入图 4.5 或图 4.6 所示的识别系统中. 类似地，当我们制作 MSF 时，我们必须采用变换后的参考信号 $f(\xi,\eta)$，而不是原来的参考信号 $f(x,y)$. 从 $f(x,y)$ 到 $f(\xi,\eta)$ 的变换，以及从 $f(ax,ay)$ 到 $f(\xi+\xi_1,\eta+\eta_1)$ 的变换，既可以用电子学的方法实现[9]，也可以借助于光学坐标变换器[11]用光学方法实现[10]. 当然最为简单而实用的办法是用变焦镜(ZOOM)，然而这将包含一个机械的动作，而且不能算作是并行处理器.

4.8 联合变换相关器

光学联合变换相关器(JTC，joint Fourier transform correlator)最早是由 Weaver 和 Goodman[12]，以及 Rau[13] 提出来的，如图 4.15 所示. 图中 L 为变换透镜. 一个具有实的透过率 $f(x,y)$ 的透明片置于输入平面一侧，其中心位于 $(-a,0)$，第二个具有实的透过率 $g(x,y)$ 的透明片置于输入平面另一侧，其中心位于 $(a,0)$. 两个透明片用准直的激光束照明，并通过透镜进行傅里

叶变换. 透镜后焦面的振幅分布为

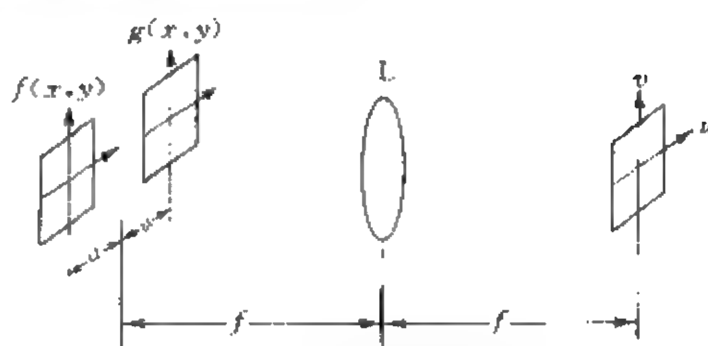


图 4.15 联合变换功率谱的记录

$$S(u,v)=\iint_{-\infty}^{\infty}[f(x+a,y)+g(x-a,y)]$$

$$\times\exp\left[-\mathrm{i}\frac{2\pi}{\lambda f}(xu-yv)\right]\mathrm{d}x\mathrm{d}y. \tag{1}$$

$S(u,v)$ 称为 $f(x,y)$ 和 $g(x,y)$ 的联合傅里叶谱,可借助于节 1.1 中关于位移的变换法则改写为

$$S(u,v)=\exp\left[\mathrm{i}\frac{2\pi}{\lambda f}au\right]F(u,v)+\exp\left[-\mathrm{i}\frac{2\pi}{\lambda f}au\right]G(u,v), \tag{2}$$

式中

$$F(u,v)=\iint_{-\infty}^{\infty}f(x,y)\exp\left[-\mathrm{i}\frac{2\pi}{\lambda f}(xu+yv)\right]\mathrm{d}x\mathrm{d}y, \tag{3}$$

$$G(u,v)=\iint_{-\infty}^{\infty}g(x,y)\exp\left[-\mathrm{i}\frac{2\pi}{\lambda f}(xu+yv)\right]\mathrm{d}x\mathrm{d}y. \tag{4}$$

如果用平方律探测器测量透镜后焦面的图形,得到光强分布

$$|S(u,v)|^2=|F(u,v)|^2+\exp\left[\mathrm{i}\frac{4\pi}{\lambda f}au\right]F(u,v)G^*(u,v)$$

$$+\exp\left[-\mathrm{i}\frac{4\pi}{\lambda f}au\right]F^*(u,v)G(u,v)+|G(u,v)|^2. \tag{5}$$

它就是联合傅里叶变换的功率谱(参见图 4.16,图中 L 为变换透镜).

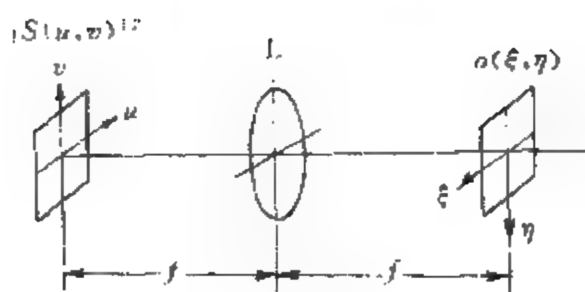


图 4.16　联合功率谱的傅里叶变换

对上述联合功率谱 $|S(u,v)|^2$ 再进行一次傅里叶逆变换，得到

$$o(\xi,\eta)=\iint_{-\infty}^{\infty}|S(u,v)|^2\exp\left[\mathrm{i}\,\frac{2\pi}{\lambda f}(\xi u+\eta v)\right]\mathrm{d}u\mathrm{d}v. \qquad (6)$$

将(5)式代入(6)式得到四项，$o=o_1-o_2+o_3+o_4$. 让我们逐项进行分析. 第一项为

$$o_1(\xi,\eta)=\iint_{\infty}^{\infty}|F(u,v)|^2\exp\left[\mathrm{i}\,\frac{2\pi}{\lambda f}(\xi u+\eta v)\right]\mathrm{d}u\mathrm{d}v, \qquad (7)$$

它是 $f(x,y)$ 的自相关，

$$o_1(\xi,\eta)=\iint_{-\infty}^{\infty}f(\alpha,\beta)f^*(\alpha-\xi,\beta-\eta)\mathrm{d}\alpha\mathrm{d}\beta. \qquad (8)$$

第二项为

$$o_2(\xi,\eta)=\iint_{\infty}^{\infty}\exp\left[\mathrm{i}\,\frac{4\pi}{\lambda f}au\right]F(u,v)G^*(u,v)$$
$$\times\exp\left[\mathrm{i}\,\frac{2\pi}{\lambda f}(\xi u+\eta v)\right]\mathrm{d}u\mathrm{d}v. \qquad (9)$$

它可以表为

$$o_2(\xi,\eta)=\iint_{-\infty}^{\infty}f(\alpha,\beta)g^*[\alpha-(\xi+2a),\beta-\eta]\mathrm{d}\alpha\mathrm{d}\beta. \qquad (10)$$

正是 $f(x,y)$ 和 $g(x,y)$ 的相关，只是在 ξ 轴上平移 $-2a$. 类似地，第三项正是 $g(x,y)$ 和 $f(x,y)$ 的相关，在 ξ 轴上平移 $2a$ 的距离：

$$o_3(\xi,\eta)=\iint_{-\infty}^{\infty}g(\alpha,\beta)f^*[\alpha-(\xi-2a),\beta-\eta]\mathrm{d}\alpha\mathrm{d}\beta. \qquad (11)$$

第四项

$$o_1(\xi,\eta) = \iint_{-\infty}^{\infty} g(\alpha,\beta) g^*(\alpha - \xi, \beta - \eta) \mathrm{d}\alpha \mathrm{d}\beta, \qquad (12)$$

为 $g(x,y)$ 的自相关. 它与第一项一起重叠在输出平面中心附近,可以称之为 0 级项,它们不是信号. 两个互相关项 o_2 和 o_3 为一级项,正是我们寻求的相关输出信号,它们在输出平面上沿 ξ 轴分别平移 $-2a$ 和 $2a$,因而与 0 级项分离. 在 f 和 g 相同的情况下,o_2 和 o_3 的函数形式相同.

4.9 实时联合变换相关器

联合变换谱的记录和逆变换两个过程之间,有一个用平方律探测器把联合变换的复振幅谱转换为功率谱的中介过程. 在早期的实验中,这一过程借助于感光胶片实现: 首先用感光胶片记录一对目标图形 $f(x,y)$ 和参考图形 $g(x,y)$ 的联合变换谱,经过显影、定影处理后,胶片的透过率近似正比于联合变换的功率谱 $|S(u,v)|^2$, 再把它放在系统的输入平面上,用透镜进行逆变换,获得相关输出. 如果要对多个目标图形 $f_n(x,y)$ $(n = 1,2,\cdots,N)$ 进行识别,上述过程必须重复多次,显然是十分费时费事的. 而在 Vander Lugt 相关器中,可以只用一个匹配滤波器来依次对多个输入图形进行识别.

然而,如果我们采用高分辨率光寻址的空间光调制器来代替胶片完成上述中介过程,联合变换相关器 JTC 的优越性就体现出来了.

早在 1981 年,Pichon 和 Huignard[14] 第一次用光折变晶体 BSO 来实时记录联合变换的功率谱. 1985 年,Loiseaux 等[15] 进一步把 BSO 放在傅里叶平面上,并在输入平面上放置 BSO 液晶光阀 (LCLV). 尽管 BSO 晶体的分辨率高达 1000 lp/mm(线对每毫米),BSO-LCLV 系统只能实现 12 lp/mm 的分辨率,可用面积为 25 mm × 25 mm,相当于 300 × 300 个像素. 这一方案实现了 JTC

中的实时平方律转换.

1984 年,Yu 和 Lu[16] 提出用 LCLV 作为平方律探测器,而在输入平面上采用计算机控制的磁光空间光调制器(MOSLM) 的实时可编程 JTC 方案,当时 MOSLM 的空间带宽积(参见第一章) 或像素数仅 48 × 48. 一般来讲,物体的形状要比它的 JTC 功率谱规矩得多,所以输入器件的分辨率可以比谱面器件的分辨率低. 由于 LCLV 的分辨率仅 15 lp/mm,比胶片或 BSO 的分辨率低,所以在上述系统中都用长焦距变换透镜,因为傅里叶谱的线度是正比于焦距的. 上述工作为实时可编程 JTC 引入了全新的观念.

到 1987 年,Yu 等又提出了另一个新方案,用液晶电视(LCTV) 作为电寻址 SLM,既用在输入面,又用在谱面. 在该方案中,计算机控制的 LCTV 并排显示 $f(x,y)$ 和 $g(x,y)$,并用 CCD 作为平方律探测器件记录输入信号的联合变换功率谱. 探测到的图形再通过 LCTV 显示出来,并由 CCD 探测相关输出. CCD 的探测面积为 10 mm × 10 mm,包含 512 × 512 个像素,这样一来分辨率达到 51.2 lp/mm. 目前液晶显示器件(LCD) 的像素数已超过 1000 × 1000,然而当时 LCTV 仅 120 × 240 个像素,因此上述 CCD-LCTV 组件的有效分辨率只有 12 lp/mm. 尽管有效分辨率与 LCTV 的分辨率的数量级相同,在这一新的系统中还是使用了附加的放大物镜. Yu 等的工作有两个特点:(1) 不采用长焦距变换透镜,因而 JTC 是相当紧凑的. (2) 用 CCD 来探测功率谱,从而在第二次傅里叶变换前可以用各种数字处理技术来提高系统的性能.

设 $f(x,y)$ 与 $g(x,y)$ 相同,则联合变换谱可以认为是 $f(x,y)$ 和 $g(x,y)$ 上对应的无数点对形成的杨氏条纹的相干叠加[18]. 杨氏条纹的衍射图包含两个亮斑(+ 1 级和 − 1 级衍射) 及 0 级(直流) 光斑. 这两个亮斑即相关峰.

现在让我们来分析 $|S(u,v)|^2$. 当 $f(x,y) = g(x,y)$ 时,节 4.8(5) 式成为

$$
\begin{aligned}
|S(u,v)|^2 = & |F(u,v)|^2 + \exp[\mathrm{i}4\pi au/\lambda f]|F(u,v)|^2 \\
& + \exp[-\mathrm{i}4\pi au/\lambda f]|F(u,v)|^2 + |F(u,v)|^2 \\
= & 2|F(u,v)|^2[1 + \cos(4\pi au/\lambda f)].
\end{aligned} \tag{1}
$$

因子$[1 + \cos(4\pi au/\lambda f)]$代表理想的杨氏条纹，周期为$(\lambda f/2a)$，而因子$|F(u,v)|^2$则是杨氏条纹的包络. 根据节 1.1 中关于卷积的变换法则，$|S(u,v)|^2$的傅里叶逆变换将是$2|F(u,v)|^2$的傅里叶逆变换与$[1 - \cos(4\pi au/\lambda f)]$的傅里叶逆变换的卷积.

$$
\begin{aligned}
o(\xi,\eta) = & \iint_{-\infty}^{\infty} |S(u,v)|^2 \exp[\mathrm{i}2\pi(\xi u + \eta v)/\lambda f]\mathrm{d}u\mathrm{d}v \\
= & \mathscr{F}^{-1}\{2|F(u,v)|^2\} * \mathscr{F}^{-1}\{1 + \cos(4\pi au/\lambda f)\},
\end{aligned} \tag{2}
$$

其中后一项经运算后化为

$$
\begin{aligned}
& \mathscr{F}^{-1}\{1 + \cos(4\pi au/\lambda f)\} \\
& = \delta(\xi,\eta) + \frac{1}{2}[\delta(\xi + 2a,\eta) + \delta(\xi - 2a,\eta)].
\end{aligned} \tag{3}
$$

我们已经知道$|F(u,v)|^2$的傅里叶逆变换是$f(x,y)$的自相关，记为$c_{ff}(\xi,\eta)$. 亦即

$$
\mathscr{F}^{-1}\{2|F(u,v)|^2\} = 2c_{ff}(\xi,\eta). \tag{4}
$$

把(3)、(4)式代入(2)得到

$$
\begin{aligned}
o(\xi,\eta) = & 2c_{ff}(\xi,\eta) * \left\{\delta(\xi,\eta) + \frac{1}{2}[\delta(\xi+2a,\eta) \right. \\
& \left. + \delta(\xi-2a,\eta)]\right\} \\
= & 2c_{ff}(\xi,\eta) + c_{ff}(\xi+2a,\eta) + c_{ff}(\xi-2a,\eta).
\end{aligned} \tag{5}
$$

从而自相关斑分别出现在$(-2a,0)$，$(0,0)$和$(-2a,0)$

杨氏条纹$[1 + \cos(4\pi au/\lambda f)]$比包络$[2|F(u,v)|^2]$重要得多，即使包络有所改变，我们仍能观察到相关峰$\delta(\xi + 2a,\eta) + \delta(\xi - 2a,\eta)$. 例如功率谱的范围超过 CCD 的探测面，仅一部分功率谱被检测[17]. 设 CCD 的探测面由矩形函数$\mathrm{rect}(u,v)$所限定($\mathrm{rect}(u,v)$是函数$\mathrm{rect}(u)\mathrm{rect}(v)$的简写). 设它比$F(u,v)$要小，则 CCD 测得的谱为$\mathrm{rect}(u,v)[1 + \cos(4\pi au/\lambda f)]$，在透镜后焦面上的输出为

$$o(\xi,\eta)=\iint_{-\infty}^{\infty}\operatorname{rect}(u,v)\left[1-\cos\left(\frac{4\pi au}{\lambda f}\right)\right]$$

$$\times\exp\left[\mathrm{i}\frac{2\pi}{\lambda f}(\xi u+\eta v)\right]\mathrm{d}u\mathrm{d}v$$

$$=c_{ss}(\xi,\eta)+\frac{1}{2}c_{ss}(\xi+2a,\eta)+\frac{1}{2}c_{ss}(\xi-2a,\eta).\quad(6)$$

在$(-2a,0)$及$(2a,0)$仍然看到相关峰(设$f(x,y)=g(x,y)$),只是相关峰的包络不是$c_{ff}(\xi,\eta)$而是$c_{ss}(\xi,\eta)$,它是$\operatorname{sinc}(x,y)$的自相关($\operatorname{sinc}(x,y)=\operatorname{sinc}(x)\operatorname{sinc}(y)$). 注意$\operatorname{sinc}(x,y)$和$\operatorname{rect}(u,v)$是傅里叶变换对,且$\operatorname{rect}(u,v)=|\operatorname{rect}(u,v)|^2$.

上述JTC具有自适应的特征,它的优点是:(1)与Vander Lugt相关器的全息匹配滤波器相比,JTC的参考图形显然要简单得多,它可以存储在计算机中,并用LCTV显示.(2)探测联合变换功率谱可以是一个非线性的处理过程.(3)由于我们对相关峰的形状并不特别感兴趣,因而可以选择功率谱的一部分来产生杨氏条纹.(4)如果采用短焦距傅里叶变换透镜加一个放大透镜,系统总长度可大大缩短.

JTC的缺点在于功率谱受到输入平面上所有信号和噪声的影响,它强烈地依赖信号和噪声的特性,很难设计一个非线性函数来优化功率谱的探测. 本节涉及的各种空间光调制器SLM的深入讨论请参见第七章,有关光折变晶体的内容请参见第九章.

4.10 联合变换相关器的应用

JTC最通常的应用是图形识别,这早在Ran的论文中就提出来了. 最近的论文报道了JTC在实时自适应图形识别中的许多应用,采用的方案都是以LCTV作为输入的空间光调制器SLM[17]. 图4.17表示一个初步的实验装置[17],图中L_1,L_2为扩束透镜;L_3,L_4为成像透镜.

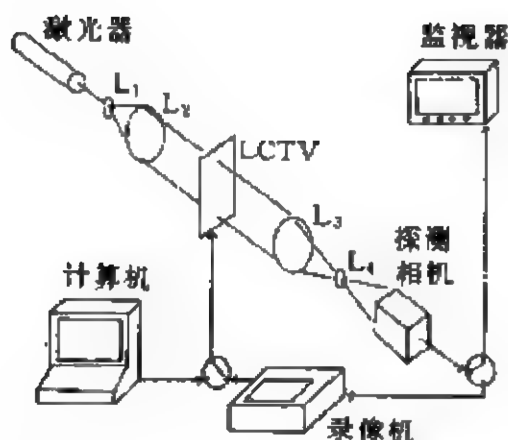


图 4.17　使用 LCTV 的实时自适应 JTC 的示意图[17]

一个微机在 LCTV 上同时产生待识别物体和参考图形. 准直的相干光入射到 LCTV 上,物体和参考图形由透镜 L_3 进行联合变换,联合变换的功率谱显示出条纹结构,由 CCD 探测并记录在磁带上,使用同一个 LCTV 显示条纹结构,如果有相关峰出现,它就被 CCD 探测并在显示终端显示出来. 如果使用高速 CCD,并用带有图像采集卡的计算机代替磁带机,就能实现高性能的实时操作.

计算机生成的物体和参考图像见图 4.18(a),(b) 中的亮点即相关输出,右方为直流光斑. 当 LCTV 不显示物体时,相关斑消失,见图(c). 图(b) 和(c) 的差则为相关输出,我们注意相减运算消除了直流光斑[17]. 当模拟物体移动时,相关斑也相应移动,说明 JTC 具有平移不变性.

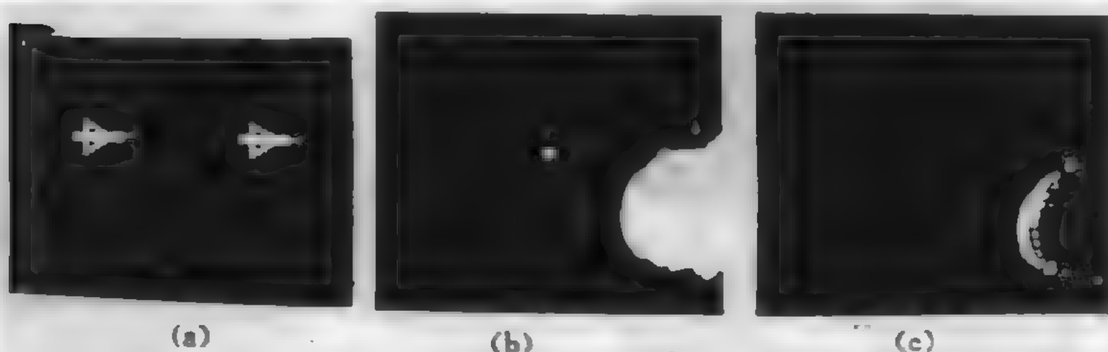


图 4.18　(a) 计算机生成的物体和参考图像　(b) 相关输出
(c) LCTV 上没有物体存在时的直流输出

Fielding 等[19] 报道了在 Yu 等[17] 方案的基础上加以改进获得的指纹识别的结果，实验装置见图 4.19.

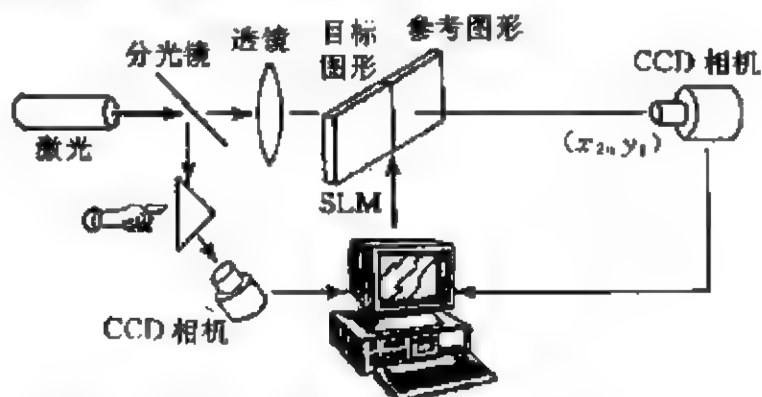


图 4.19 JTC 指纹识别仪示意图

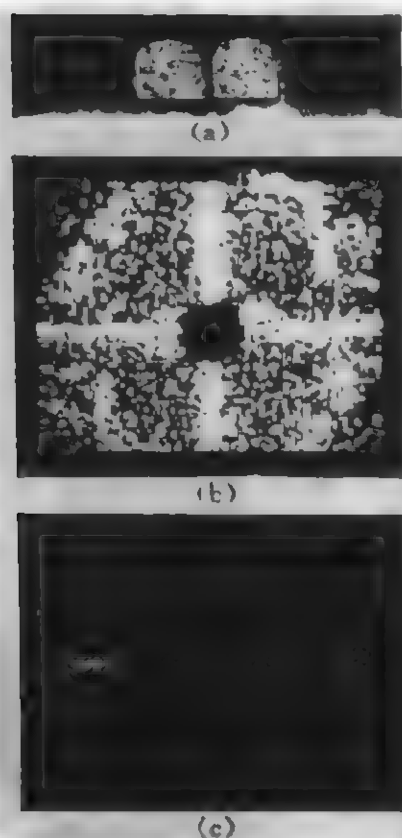


图 4.20 运用 JTC 进行指纹识别的实验结果[19]
(a) 第一步，输入指纹和参考指纹经二值化处理后在输入平面显示
(b) 第二步，二值化的联合变换功率谱
(c) 一对相关斑

指纹由 CCD 观察，并送入计算机，通过计算机将该指纹与参考指纹在空间光调制器 SLM 上同时显示. 第二个 CCD 探测联合变换功率谱，功率谱信号经计算机进行二值化处理后，通过该 SLM 显示出来. 最后产生的相关输出由 CCD 探测，实验结果见图 4.20. 注意图 (b)、(c) 的中心部分已被挡去.

宋菲君等[20] 利用实时 JTC 相关器成功地识别了破损的指纹. 图 4.21(a) 为参考指纹，(b) 是从犯罪现场获得的破损的指纹，(c) 是相关输出，一对相关亮斑仍清晰可见. 图 4.22(a) 是另一个参考指纹，(b) 是罪犯留在电子手表面上经处理后的现场指纹，(c)是

它们的相关输出，相关峰相当清晰.

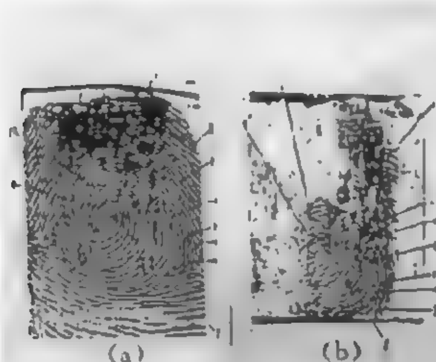


图 4.21 (a) 参考指纹
(b) 现场指纹 (c) 相关输出[20]

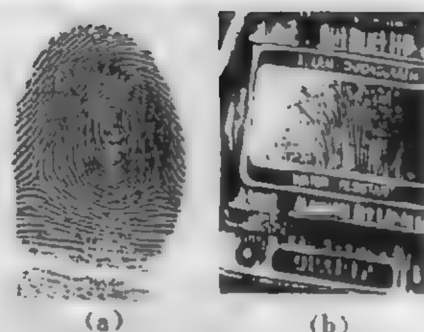


图 4.22 (a) 参考指纹
(b) 留在电子表表面上的现场指纹(经处理)
(c) 相关输出[20]

由 Yu 等[17] 提出来的 JTC 方案的主要优点就在于功率谱由 CCD 探测并转变为电信号，这样一来，在送到 LCTV 显示并进行第二次傅里叶变换前，可以先用计算机对联合变换功率谱进行预处理，例如二值化的处理[19]，许多方案都采用这一类光电混合处理并获得满意的结果.

采用杨氏条纹的概念对 JTC 的原理进行证明[18]，使我们联想到二次曝光散斑计量术[21]，两者非常相似. 在二次曝光散斑计量术中，首先记录物体的散斑图. 当物体位移后，散斑场没有发生形状分布的变化，只发生位移，散斑图的位移和物体的位移成正比. 在同一张照相底片上记录位移后的散斑图，经处理后的底片上包

含两个相同的散斑场图案.底片的衍射图系两个散斑图中所有相应的点对产生的杨氏条纹的叠加,再把杨氏条纹记录在一张底片上.杨氏条纹的衍射图包含一对分布在中心左右$(-d,0)$及$(d,0)$处的光斑,由此可以确定物体发生的位移.在天文学中,用这一方法来分辨双星[22].现在已能实现实时二次曝光散斑术[22],使用的仍然是Yu等提出的JTC方案[12],实验结果见图4.23.

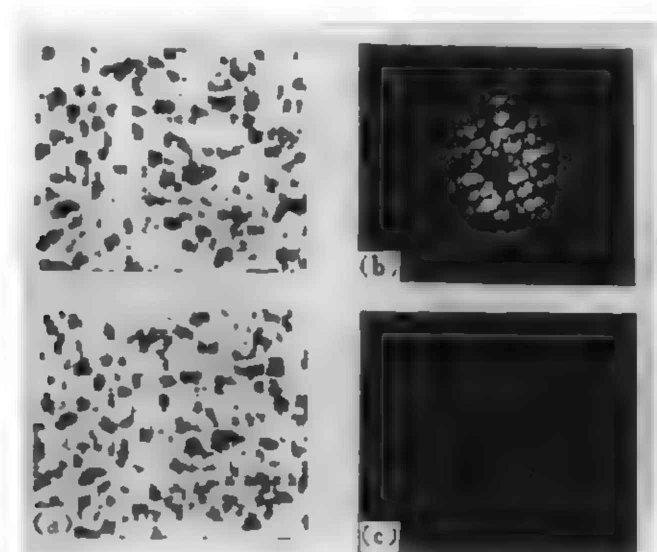


图4.23 实时二次曝光散斑计量术的实验结果[22]
(a) 二次曝光位移散斑图 (b) 联合变换功率谱 (c) 相关输出

4.11 旋转不变联合变换相关器

在机器人视觉的各种应用中,旋转不变图形识别是至关重要的.工业机器人的使用环境比军事运用要好一些.如果把照相机与实物平面之间的距离固定下来,那么相关器对于物体比例的敏感性就不那么重要了,在许多实际加工厂的图像识别应用中旋转不变就足够了,例如,对一条传送带上的某种部件的识别.

圆谐函数展开是一种最为直截了当的实现旋转不变的方案[8],在Vander Lugt相关器[8]、JTC[24]及非相干相关器[25]中都已得到成功.然而这一方法的一个根本问题在于,旋转不变的性能在

很大程度上取决于旋转中心的选择，而选择旋转中心并没有固定的法则可以遵循. 此外，还需要一台计算机来产生滤波器. 这里我们来看另一个途径，即旋转参考图形[26]，这一方法简单而有趣.

在 JTC 中，物体函数 $f(x,y)$ 和参考图形函数 $g(x,y)$ 并排显示在输入平面上，在傅里叶平面上产生联合傅里叶变换谱. 使用空间光调制器 SLM 等平方律探测器，SLM 上的调制光强的幅度就正比于 $f(x,y)$ 和 $g(x,y)$ 的联合傅里叶变换功率谱，进一步对联合功率谱的傅里叶变换产生了 $f(x,y)$ 和 $g(x,y)$ 的相关.

为简单起见，设参考图形 $g(x,y)$ 为字母“A”，输入图形则包含数个字母“A”，散布在其他字母中，见图 4.24(a). 如果将 $g(x,y)$ 从 0° 旋转到 360°，每 1° 取一个相关，对所有的相关求和，我们就会得到一幅旋转不变的相关输出，如图 4.24(b) 所示. (a) 中每一个字母“A”，不管它取向如何，都在(b) 的相应位置产生一个相关斑，而所有其他字母则不生成任何输出.

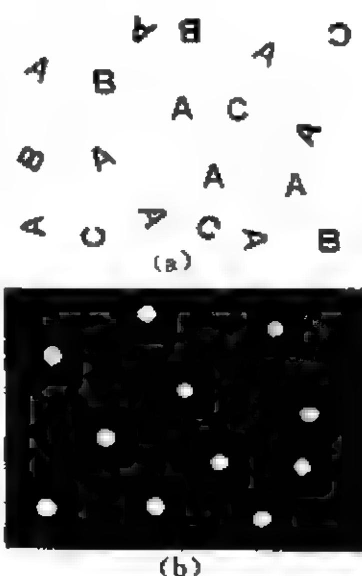


图 4.24 (a) 输入图形中包含一系列各种取向的字母“A” (b) 不论取向如何，每个字母“A”产生一个相关斑

显然不必旋转输入图形 $f(x,y)$，倘若我们旋转 $f(x,y)$，势必造成输入平面上字母“A”位置的混乱，难以恢复. 旋转已知的参考图形 $g(x,y)$ 相对来说要简单得多，旋转中心即字母“A”的几何中心，因而在旋转过程中图形几乎不发生位移，如果旋转中心与几何中心不一致，只要相距不远，旋转过程中字母的位移量也不会很大.

上述系统的示意图见图 4.25. 我们采用了改进后的 JTC 的架构[27]，在此架构中使用分光镜 BS，把准直的激光束分成两束，经 M_1 和 M_2 两个反射镜反射，一个光束通过可旋转的参考图形 $g(x,y)$，另一个则通过液晶电视(LCTV)，由 CCD_1 提供了输入图形

$f(x,y)$. 两个调制后的光束再经过棱镜BS_2合成，通过透镜FT_1进行傅里叶变换，在液晶光阀LCLV的输入端面形成联合变换谱，这一信号经LCLV转换成联合变换功率谱，由另一准直的激光束通过偏振分光棱镜PBS读出，并经第二个透镜FT_2进行第二次傅里叶变换，产生相关输出，由CCD_2探测. 参考图形$g(x,y)$放置在可旋转的中空转台上，转台由步进电机驱动旋转.

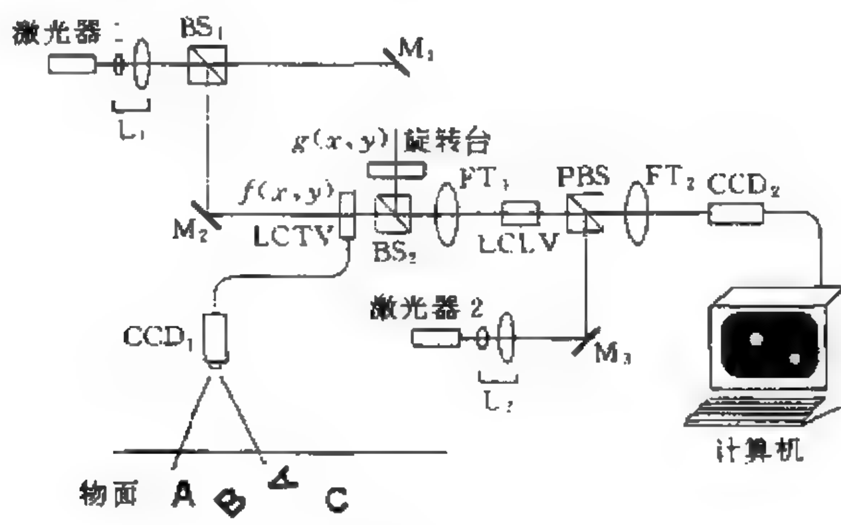


图 4.25

将联合变换谱写入LCLV的激光束能量为24×10^{-7} J/cm²，LCLV开启和关闭的时间分别为10 ms和15 ms，即LCLV把一幅图像写入和读出的时间至少为25 ms. 设每1°执行一次LCLV的存取操作，则完成一个完整的旋转不变识别操作共需360×25 ms $= 9$ s的时间. CCD_2工作在积分模式，积分时间是9 s. 在这个颇长的时间内，要求环境是暗的，亦即除相关输出信号外，环境光不能干扰CCD_2. 我们也可以使CCD_2工作在普通的模式下，并使参考信号的旋转与CCD_2的帧频(30帧/s)同步，并在计算机中用数字方法把各帧的图像相加. 这种模式下，旋转不变相关识别操作共需12 s.

事实上，旋转参考图形并不是新的想法，早在Vander Lugt相

关识别中，Fujii 等[28] 就曾提出这种方法，他们把全息空间匹配滤波器(SMF) 放在精密转台上试图实现旋转不变相关. 然而 SMF 的中心必须与转台的中心严格重合，两者稍有偏离，即引起相关峰光强迅速下降，因此实验是很难做的. 然而在 JTC 中，参考图形“A” 的几何中心与转台中的不必严格重合，两者的偏差不会影响相关峰的强度，只是使峰的位置发生位移，位移量与中心的偏差成正比例.

4.12 结　论

本章介绍了非相干光相关器. Vander Lugt 相关器(VLC) 和联合傅里叶变换相关器(JTC). Vander Lugt 相关和联合相关都通过对两个函数傅里叶变换的积进行傅里叶逆变换的方案实现相关. 非相干相关器输出的反差太小，因而还不实用. JTC 与 VLC 相比，具有容易进行实时识别以及自适应的特点. 如果输入中包含多个物体时，JTC 的性能迅速变差，然而当匹配空间滤波器制作成功后，输入中包含物体的个数对 VLC 的性能不产生影响. 本章还介绍了光学相关识别器的许多应用，这一领域的研究至今仍在进行.

参　考　文　献

[1] S. Jutamulia, Ed., Selected Papers on Optical Correlators, SPIE Milestone Series, MS 76, SPIE, Bellingham, Washington, 1993.

[2] D. McLachlan, Jr., The role of optics in applying correlation functions to pattern recognition, *Jour. Opt. Soc. Am.* **52** (1962), 454-459.

[3] A. Vander Lugt, Signal detection by complex spatial filtering, *IEEE Trans. Inf. Theory*, **IT-10** (1964), 139-145.

[4] A. D. Gara, Real-time tracking of moving objects by optical correlation, *Appl. Opt.*, **18**(1979), 172-174.

[5] J. G. Duthie and J. Upatnieks, Compact real-time coherent optical correlator, *Opt. Eng.*, **23** (1984), 7-11.

[6] D. A. Gregory, J. C. Kirsch, J. A. Loudin, Optical correlators, optical

computing that really works, *Proc. SPIE*, **1296**(1990), 2-19.

[7] F. T. Gamble, L. M. Frye, D. R. Grieser, Real-time fingerprint verification system, *Appl. Opt.*, **31**(1991), 652-655.

[8] Y.-N. Hsu and H. H. Arsenault, Optical pattern recognition using circular harmonic expansion, *Appl. Opt.*, **21** (1982), 4016-4019.

[9] D. Casasent and D. Psaltis, Scale invariant optical transform, *Opt. Eng.*, **15**(1976), 258-261.

[10] Y. Saito, S. Komatsu, H. Ohzu, Scale and rotation invariant real time optical correlator using computer generated hologram, *Opt. Commun.*, **47**(1983), 8-11.

[11] O. Bryngdahl, Geometrical transformations in optics, *Jour. Opt. Soc. Am.*, **64**(1974), 1092-1099.

[12] C. S. Weaver and J. W. Goodman, A technique for optically convolving two functions, *Appl. Opt.*, **5**(1966), 1248-1249.

[13] J. E. Rau, Detection of differences in real distributions, *J. Opt. Soc. Am.*, **56**(1966), 1490-1494.

[14] L. Pichon and J. P. Huignard, Dynamic joint-Fourier transform correlator by Bragg diffraction in photorefractive $Bi_{12}SiO_{20}$ crystals, *Opt. Commun.*, **36** (1981), 277-280.

[15] B. Loiseaux, G. Illiaquer, and J. P. Huignard, Dynamic optical cross-correlator using a liquid crystal light valve and a bismuth silicon oxide crystal in the Fourier plane, *Opt. Eng.*, **24** (1985), 144-149.

[16] F. T. S. Yu and X. J. Lu, A programmable joint transform correlator, *Opt. Commun.*, **52**(1984), 10-16.

[17] F. T. S. Yu, S. Jutamulia, T. W. Lin, and D. A. Gregory, Adaptive real-time pattern recognition using a liquid crystal TV based joint transform correlator, *Appl. Opt.*, **26**(1987), 1370-1372.

[18] S. Jutamulia, Joint transform correlators and their applications, *Proc. SPIE*, **1812**(1992), 233-243.

[19] K. H. Fielding, J. L. Horner, and C. K. Makekau, Optical fingerprint identification by binary joint transform correlation, *Opt. Eng.*, **30**(1991), 1958-1961.

[20] F. Song, X. An, J. Song, P. Ji, C. Bai, T. Lin, B. Liu, H. Huang, Fingerprint identification by using a joint-Fourier correlator and a liquid crystal light volve, *Opt. Laser Tech.*, **26** (1994), 29-37.

[21] E. Archbold, J. M. Burch, and A. E. Ennos, Recording of in-plane surface displacement by double exposure speckle photography, *Opt. Acta*, **17**(1970), 883.

[22] A. Labeyrie, Attainment of diffraction limited resolution in large telescopes by Fourier analysing speckle pattern in star images, *Astron. & Astrophys.*, **6** (1970), 85.

[23] A. Ogiwara, H. Sakai, and J. Ohtsubo, Real-time optical joint transform correlator for velocity measurement using clipped speckle intensity, *Opt. Commun.*, **78** (1990), 322-326.

[24] F. T. S. Yu, X. Y. Li, E. Tam, S. Jutamulia, and D. A. Gregory, Rotation-invariant pattern recognition with a programmable joint transform correlator, *Appl. Opt.*, **28** (1989), 4725-4727.

[25] E. Elizur and A. A. Friesem, Rotation-invariant correlation with incoherent light, *Appl. Opt.*, **30**(1991), 4175-4178.

[26] S. Jutamulia and T. Asakura, Rotation-invariant joint transform correlator, *Appl. Opt.*, **33**(1994), 5440-5442.

[27] S. Jutamulia and D. A. Gregory, Photo-compact-disk-based optical correlator, *Appl. Opt.*, **32**(1993), 5092-5094.

[28] H. Fujii, S. P. Almeida, and J. E. Dowling, Rotational matched spatial filter for biological pattern recognition, *Appl. Opt.*, **19**(1980), 1190-1195.

第五章　广义傅里叶变换及其光学实现

5.1 引　　言

从第一章的讨论中我们知道，形如

$$\Psi(u,v) = \mathscr{F}\{\psi_0\} = \iint_{-\infty}^{\infty} \psi_0(x,y)\exp[-\mathrm{i}2\pi(ux + vy)]\mathrm{d}x\mathrm{d}y \tag{1}$$

的二维傅里叶变换可以用光学系统近似实现. 当用单色光照射位于薄透镜或透镜系统前焦面的二维图像 $\psi_0(x,y)$ 时，在后焦面上将出现它的傅里叶变换. 在本章中我们将进一步研究当物体到透镜的距离 d_1 及输出图像到透镜的距离 d_2 不等于透镜的焦距 f 时透镜或透镜系统对输入图像的变换.

研究表明，d_1 和 d_2 满足一定的条件时，输出平面上将出现 ψ_0 的广义傅里叶变换：

$$\mathscr{F}_\alpha\{\psi_0\} = \iint_{-\infty}^{\infty} \psi_0(x,y)\exp\left[\mathrm{i}\pi\frac{(x^2+u^2)+(y^2+v^2)}{\lambda\tilde{f}\tan\alpha} - \mathrm{i}2\pi\frac{(xu+yv)}{\lambda\tilde{f}\sin\alpha}\right]\mathrm{d}x\mathrm{d}y, \tag{2}$$

又称为分数阶傅里叶变换(fractional Fourier transform)，它显然是常规傅里叶变换(1)式的推广，当 $\alpha=\pi/2$ 时就变成(1)式.

早在 1937 年，Condon 在他的论文[1]中就提出了广义傅里叶变换的初步概念. 1961 年，Bargmann[2]又进一步发展了这些概念，他认为变换可以分别用厄米多项式和积分变换加以定义. 而两种定义实质上是相同的.

到 1980 年，Namias[3]完整地提出了广义傅里叶变换的数学定义、性质，讨论了变换的本征函数，并用于处理谐振子的薛定谔方程、格林函数问题、在均匀磁场中的自由电子的能级、在含时间变量的均匀磁场中自由电子薛定谔方程的求解等. 1987 年，McBride 和 Kerr[4]进一步研究了广义傅里叶变换，把变换看作是充分光滑的函数构成的向量空间(Frechet 空间)中的算子，在此框架内建立了广义傅里叶变换更为严谨、完整的理论系统，这两篇文章至今仍是广义傅里叶变换的理论基础.

在 80 年代对于广义傅里叶变换的研究似乎显得很沉寂，很少文章涉及这一专题，这种情况一直持续到 90 年代，光学界的科学家和工程师开始关注广义傅里叶变换与光学的关系，与三十年前常规傅里叶变换与光学的结合产生了傅里叶光学的情况非常相似.

1993 年，Ozaktas 和 Mendlovic[5,6]提出用平方折射率光波导(GRIN)来实现广义傅里叶变换；Lohmann[7,9]，Bernardo 等[8]则用透镜系统成功地实现了这一变换；Lohmann[9]还设计了阶数连续可变的广义光学傅里叶变换系统；Bernardo 等认为应正确地称这一变换为广义傅里叶变换，而不是分数阶傅里叶变换，因为阶数既可以是整数、分数，还可以是复数，他们还初步讨论了虚数阶傅里叶变换的概念. 我们认为称此变换为广义傅里叶变换更为合理. 从此，广义光学傅里叶变换作为为数学和光学的一个交叉领域，变得异常活跃起来.

Lohmann[9]，Mendlovic[10]阐明了广义傅里叶变换与维格纳变换的关系，指出可以用维格纳空间中的旋转来一般地定义广义傅里叶变换，这一定义与光波在梯度折射率介质中的传播的定义是等价的.

Mendlovic 等进一步讨论用广义傅里叶变换来表征信号的新方法[11]，以及分数阶光学相关[12]；Dorsch[13]，Bernardo 等[14,17]分别提出了用光学系统实现任意阶傅里叶变换的方案；Ozaktas[15]

等研究了广义傅里叶变换与小波变换的关系，他们认为广义傅里叶变换可以表为小波变换，小波函数具有 $h(x)=\exp(\mathrm{i}\pi x^2)$ 的形式．然而该函数是分布在 $(-\infty,\infty)$ 上的振荡函数，并不具备小波的特点．易证 $h(x)$ 的傅里叶变换 $H(u)=\sqrt{\mathrm{i}}\exp(\mathrm{i}\pi u^2)$，而 $H(0)\neq 0$，不符合小波变换的相容性条件(参见节 6.3)．因而我们认为广义傅里叶变换只是形式上与小波变换相似．

Mendlovic 等[16]对变换的形式稍加改换，定义了广义余弦变换，该变换适用于非相干光，在数字成像、非相干光信息处理方面都有潜在的应用．众所周知，夫琅和费衍射可以实现常规的傅里叶变换，Pellat-Finet[18]则探讨了菲涅耳衍射与广义傅里叶变换的关系．

傅里叶变换在科学技术的许多领域中有广泛的应用，因此我们可以预料广义傅里叶变换的应用领域将更为宽广．目前，它已成为数学、量子力学中重要的应用工具．

本章将研究广义傅里叶变换的数学定义、性质及实现广义傅里叶变换的光学系统，并讨论与广义傅里叶变换有密切关系的维格纳变换．

5.2 广义傅里叶变换的定义及性质

5.2.1 广义傅里叶变换的定义

为简单起见，我们仅讨论一维函数的广义傅里叶变换，有关的定义和性质可以直接推广到二维的情况．

函数 $g(\xi)$ 的广义傅里叶变换定义为

$$\mathscr{F}_\alpha\{g(\xi)\}=\left\{\frac{\exp[-\mathrm{i}(\pi/2-\alpha)]}{2\pi\sin\alpha}\right\}^{1/2}\exp\left(\frac{\mathrm{i}x^2}{2\tan\alpha}\right)\times\int_{-\infty}^{\infty}\exp\left(\frac{\mathrm{i}\xi^2}{2\tan\alpha}-\frac{\mathrm{i}x\xi}{\sin\alpha}\right)g(\xi)\mathrm{d}\xi\quad(|\alpha|<\pi),\tag{1}$$

通常称它为 $g(\xi)$ 的广义傅里叶谱，记为 $G(x)$．

以$-\alpha$代替上式中的α，得到

$$\mathscr{F}_{-\alpha}\{g(\xi)\}=\left\{\frac{\exp[\mathrm{i}(\pi/2-\alpha)]}{2\pi\sin\alpha}\right\}^{1/2}\exp\left(-\frac{\mathrm{i}x^2}{2\tan\alpha}\right)$$
$$\times\int_{-\infty}^{\infty}\exp\left(-\frac{\mathrm{i}\xi^2}{2\tan\alpha}+\frac{\mathrm{i}x\xi}{\sin\alpha}\right)g(\xi)\mathrm{d}\xi\quad(|\alpha|\leqslant\pi),\tag{2}$$

α称为广义傅里叶变换的阶.

下面我们来证明$\mathscr{F}_{-\alpha}$是$\mathscr{F}_{\alpha}$的逆变换：

$$\mathscr{F}_{-\alpha}\mathscr{F}_{\alpha}\{g(\xi)\}=\frac{1}{2\pi\sin\alpha}\int_{-\infty}^{\infty}\exp\left[\frac{\mathrm{i}(\xi^2-x^2)}{2\tan\alpha}\right]g(\xi)I\mathrm{d}\xi,\tag{3}$$

式中的积分

$$I=\int_{-\infty}^{\infty}\exp\left[\frac{\mathrm{i}(x-\xi)\xi'}{\sin\alpha}\right]\mathrm{d}\xi'=2\pi\sin\alpha\cdot\delta(\xi-x),$$

代入(3)式得到

$$\mathscr{F}_{-\alpha}\mathscr{F}_{\alpha}\{g(\xi)\}=g(x).\tag{4}$$

下面我们将用更加普遍的方法来证明$\mathscr{F}_{-\alpha}$是$\mathscr{F}_{\alpha}$的逆变换.

在广义傅里叶变换中，α可以取区间$(-\pi,\pi]$中的任意实数. $(-\pi,\pi]$称广义傅里叶变换的主值区间，当α超出主值区间时，相应的变换可以化成在该区间内的变换，下文中将证明这一点. 因此$\mathscr{F}_{-\alpha}(\alpha>0)$实质上只是负阶数的广义傅里叶变换，并没有特别的意义. 所以(2)式与(1)式在变换前后的宗量及函数形式一致.

广义傅里叶变换的一个重要性质，在于当$\alpha=\pi/2$以及$\alpha=-\pi/2$时化成常规的傅里叶变换及逆变换：

$$\mathscr{F}_{\pi/2}\{g(\xi)\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}g(\xi)\exp(-\mathrm{i}x\xi)\mathrm{d}\xi,\tag{5}$$

$$\mathscr{F}_{-\pi/2}\{g(\xi)\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}g(\xi)\exp(\mathrm{i}x\xi)\mathrm{d}\xi.\tag{6}$$

以上结果是在$\alpha\to\pi/2$时取极限得到的. 注意在本章中傅里叶变换的表达式与本书其他各章有所不同.

(1)式定义的变换当$\alpha=0$时，没有意义，因而$\mathscr{F}_0$也必须另行

定义.

我们来计算 $\alpha\to 0$ 时的广义傅里叶变换.

由于 $\alpha\approx 0$ 时, $\sin\alpha\approx\alpha$, $\tan\alpha\approx\alpha$, 所以有

$$\mathscr{F}_\alpha\big|_{\alpha\to 0}\approx\int_{-\infty}^{\infty}\frac{\exp[-(x-\xi)^2/(\mathrm{i}2\alpha)]}{\sqrt{\mathrm{i}2\pi\alpha}}g(\xi)\mathrm{d}\xi=g(x), \tag{7}$$

其中用到极限意义下的 δ 函数的定义:

$$\lim_{\varepsilon\to 0}\frac{\exp(-x^2/\mathrm{i}\varepsilon)}{\sqrt{\mathrm{i}\pi\varepsilon}}=\delta(x). \tag{8}$$

从而我们可以用上述极限过程来定义 $\mathscr{F}_0$:

$$\mathscr{F}_0\{g(\xi)\}=g(x). \tag{9}$$

用类似的方法还可以定义 $\mathscr{F}_\pi$:

$$\mathscr{F}_\pi\{g(\xi)\}=g(-x). \tag{10}$$

以上两式表明: 0 阶广义傅里叶变换给出输入图像本身, π 阶广义傅里叶变换则给出它的倒像.

表 5.1 给出一些常用函数的广义傅里叶变换. 式中 $H_n(x)$ 为

表 5.1 广义傅里叶变换表[3]

	$g(x)$	$G(x)=\mathscr{F}_\alpha\{g\}$
1	$\exp(-x^2/2)$	$\exp(-x^2/2)$
2	$H_n(x)\exp(-x^2/2)$	$\mathrm{e}^{-\mathrm{i}n\alpha}H_n(x)\exp(-x^2/2)$
3	$\exp(-x^2/2+ax)$	$\exp\left(-\frac{x^2}{2}+\frac{\mathrm{i}a^2\mathrm{e}^{\mathrm{i}\alpha}\sin\alpha}{2}+ax\mathrm{e}^{-\mathrm{i}\alpha}\right)$
4	$\delta(x)$	$\left\{\frac{\exp[-\mathrm{i}(\pi/2-\alpha)]}{2\pi\sin\alpha}\right\}^{1/2}\exp\left(\mathrm{i}\frac{x^2}{2}\cot\alpha\right)$
5	$\delta(x-a)$	$\left\{\frac{\exp[-\mathrm{i}(\pi/2-\alpha)]}{2\pi\sin\alpha}\right\}^{1/2}\exp\left[\mathrm{i}\frac{(x^2+a^2)\cot\alpha}{2}-\mathrm{i}xa\csc\alpha\right]$
6	1	$\left[\frac{\exp(\mathrm{i}\alpha)}{\cos\alpha}\right]^{1/2}\exp\left(-\mathrm{i}\frac{x^2}{2}\tan\alpha\right)$
7	$\exp(\mathrm{i}kx)$	$\left[\frac{\exp(\mathrm{i}\alpha)}{\cos\alpha}\right]^{1/2}\exp\left[-\mathrm{i}\frac{(x^2+k^2)}{2}\tan\alpha-\mathrm{i}kx\csc\alpha\right]$

n 阶厄米多项式. 由于 $H_0(x)=1$,所以表中第 1 项是第 2 项的特例. 我们注意到厄米多项式与高斯函数乘积的广义傅里叶变换的结果只是增加了一个因子 $e^{-in\alpha}$,对此,节 5.3 还要详细加以讨论.

在推导表 5.1 中的变换时,用到积分公式

$$\int_{-\infty}^{\infty}\exp(-p^2x+qx)\mathrm{d}x=\frac{\sqrt{\pi}}{p}\exp\left(\frac{q^2}{4p^2}\right), \tag{11}$$

及

$$\int_{-\infty}^{\infty}\exp[\mathrm{i}(A\xi^2+B\xi)]\mathrm{d}\xi=\sqrt{\frac{\mathrm{i}\pi}{A}}\exp\left(-\mathrm{i}\frac{B^2}{4A}\right), \tag{12}$$

式中 p,q,A 和 B 为常数,且有 $p\neq0,A\neq0$.

虽然广义傅里叶变换仍然是线性变换,即:

$$\mathscr{F}_\alpha\{Ag(\xi)+Bh(\xi)\}=A\mathscr{F}_\alpha\{g(\xi)\}+B\mathscr{F}_\alpha\{h(\xi)\}, \tag{13}$$

式中 A、B 为常数,但由于变换公式中出现二次相因子,所以它的性质和常规的傅里叶变换有了很大的差别. 例如,它不再满足类似节 1.1(9) 式的缩放规律.

5.2.2 基本性质和运算法则

下面我们介绍广义傅里叶变换的基本性质. 如上所述,$g(\xi)$ 的广义傅里叶变换谱函数记为 $G(x)$,并用 $g\Leftrightarrow G$ 表示变换对.

(1) 位移(shift)

$$g(\xi+a)\Leftrightarrow\exp\left[\mathrm{i}a\sin\alpha\left(x+\frac{a\cos\alpha}{2}\right)\right]G(x+a\cos\alpha), \tag{14}$$

可以看到它和常规傅里叶变换的相应公式节 1.1(11) 式已有了很大的差别,而当 $\alpha=\pi/2$ 时即化为傅里叶变换的位移公式.

(2) 宗量乘积(multification)

令

$$G(x)=\mathscr{F}_\alpha\{g\}=A_\alpha\exp(B_\alpha)\int_{-\infty}^{\infty}\exp\left(\mathrm{i}\frac{\xi^2}{2\tan\alpha}-\mathrm{i}\frac{x\xi}{\sin\alpha}\right)g(\xi)\mathrm{d}\xi, \tag{15}$$

式中

$$A_\alpha = \left\{\frac{\exp[-\mathrm{i}(\pi/2-\alpha)]}{2\pi\sin\alpha}\right\}^{1/2}, \tag{16}$$

$$B_\alpha = \mathrm{i}x^2/2\tan\alpha. \tag{17}$$

设 $\mathrm{D}=\mathrm{d}/\mathrm{d}x$ 为微分算符，将它作用于(15)式两边，交换微分与积分的次序，得到

$$\mathrm{D}G(x)=\frac{A_\alpha\exp(B_\alpha)}{\mathrm{i}\sin\alpha}\int_{-\infty}^{\infty}\exp\left(\mathrm{i}\frac{\xi^2}{2\tan\alpha}-\mathrm{i}\frac{x\xi}{\sin\alpha}\right)\xi g(\xi)\mathrm{d}\xi+G(x)\frac{\mathrm{d}B_\alpha}{\mathrm{d}x}. \tag{18}$$

从而导出

$$\mathscr{F}_\alpha\{\xi g(\xi)\} = (x\cos\alpha + \mathrm{i}\sin\alpha\ \mathrm{D})G(x). \tag{19}$$

依此类推，设 $m\geqslant0$，得到算符表达式

$$\mathscr{F}_\alpha\{\xi^m g(\xi)\} = (x\cos\alpha + \mathrm{i}\sin\alpha\ \mathrm{D})^m. \tag{20}$$

例如设 $m=2$，有

$$(x\cos\alpha + \mathrm{i}\sin\alpha\ \mathrm{D})^2 = \cos\alpha(x^2\cos\alpha + \mathrm{i}\sin\alpha) + \mathrm{i}x\sin2\alpha\ \mathrm{D} - \sin^2\alpha\ \mathrm{D}^2. \tag{21}$$

类似可以导出微分算符 D 的幂与函数的积，以及宗量与 D 的幂和函数乘积的变换式.

(3) 微分(differentiation)

$$\mathscr{F}_\alpha\{\mathrm{D}^m g(\xi)\} = (\mathrm{i}x\sin\alpha + \cos\alpha\ \mathrm{D})^m G(x). \tag{22}$$

(4)宗量-微分混合积(mixed product)

$$\mathscr{F}_\alpha\{(\xi\ \mathrm{D})^m g(\xi)\} = [-(\sin\alpha - \mathrm{i}x^2\cos\alpha)\sin\alpha + x\cos2\alpha\ \mathrm{D} + \mathrm{i}\sin\alpha\cos\alpha\ \mathrm{D}^2]^m G(x). \tag{23}$$

(5) 指数(exponential)

运用变换的定义式(3)，经简单运算直接得到

$$\mathscr{F}_\alpha\{\mathrm{e}^{\mathrm{i}b\xi}g(\xi)\} = \exp[\mathrm{i}b\cos\alpha(x - b\sin\alpha/2)]G(x - \sin\alpha). \tag{24}$$

(6) 可加性(additivity)

广义傅里叶变换应具备两个主要特征，首先当 $\alpha=\pi/2$及 $\alpha=$

$-\pi/2$时可化为常规的傅里叶变换及逆变换，这一点我们已在前面证明过了；其次它具有可加性，亦即 α 阶及 β 阶的变换依次作用的结果应相当于 $\alpha+\beta$ 阶变换，下面我们就来证明这一点，也就是算符的乘法.

利用定义，我们得到

$$\mathscr{F}_\alpha\{\mathscr{F}_\beta\{g(\xi)\}\} = \frac{1}{2\pi}\left\{\frac{\exp(-\mathrm{i}[\pi/2-(\alpha+\beta)])}{\sin\alpha\sin\beta}\right\} \times \exp(\mathrm{i}x^2/2\tan\alpha)I(x), \tag{25}$$

式中积分

$$\begin{aligned} I(x) &= \int_{-\infty}^{\infty}\exp\left[\mathrm{i}\frac{\xi^2}{2}(\cot\alpha+\cot\beta)-\mathrm{i}\frac{x\xi}{\sin\alpha}\right] \\ &\quad \times \int_{-\infty}^{\infty}\exp\left\{\mathrm{i}\frac{\xi'}{2}\cot\beta-\mathrm{i}\frac{\xi\xi'}{\sin\beta}\right\}f(\xi')\mathrm{d}\xi\mathrm{d}\xi' \\ &= \int_{-\infty}^{\infty}f(\xi')\exp\left(\frac{\mathrm{i}}{2}\cot\beta\right)\int_{-\infty}^{\infty}\exp\left[\mathrm{i}\frac{\xi^2}{2}(\cot\alpha+\cot\beta)\right. \\ &\quad \left.+\mathrm{i}\xi\left(\frac{x}{\sin\alpha}+\frac{\xi'}{\sin\beta}\right)\right]\mathrm{d}\xi\mathrm{d}\xi'. \end{aligned} \tag{26}$$

运用积分公式(12)于上式内部的积分，并将结果代入(25)式，得到

$$\begin{aligned} \mathscr{F}_\alpha\mathscr{F}_\beta\{g(\xi)\} &= \left\{\exp\left[-\mathrm{i}\left(\frac{\pi}{2}-\frac{\alpha+\beta}{2}\right)\right]\Big/2\pi\sin(\alpha+\beta)\right\}^{1/2} \\ &\quad \times \int_{-\infty}^{\infty}\exp\left\{-\mathrm{i}\frac{\xi^2}{2}\left[\frac{\sin\alpha}{\sin(\alpha+\beta)\sin\beta}-\cot\beta\right]\right. \\ &\quad -\mathrm{i}\frac{x^2}{2}\left[\frac{\sin\beta}{\sin(\alpha+\beta)\sin\alpha}-\cot\alpha\right] \\ &\quad \left.-\mathrm{i}\frac{x\xi}{\sin(\alpha+\beta)}\right\}f(\xi)\mathrm{d}\xi, \end{aligned} \tag{27}$$

上式积分号后的方括号[]内的三角函数运算结果均为 $-\cot(\alpha+\beta)$，我们最后得到

$$\mathscr{F}_\alpha\mathscr{F}_\beta\{g(\xi)\} = \left\{\exp\left[-\mathrm{i}\left(\frac{\pi}{2}-\frac{\alpha+\beta}{2}\right)\right]\Big/2\pi\sin(\alpha+\beta)\right\}^{1/2}$$

$$\times \exp\left(\frac{\mathrm{i}x^2}{2\tan(\alpha+\beta)}\right)\int_{-\infty}^{\infty}\exp\left[\frac{\mathrm{i}\xi^2}{2\tan(\alpha+\beta)} - \frac{\mathrm{i}x\xi}{\sin(\alpha+\beta)}\right]g(\xi)\mathrm{d}\xi$$

$$= \mathscr{F}_{\alpha+\beta}\{g(\xi)\}. \tag{28}$$

由于上式中 α 和 β 是对称的，所以有

$$\mathscr{F}_{\alpha}\mathscr{F}_{\beta}\{g\} = \mathscr{F}_{\beta}\mathscr{F}_{\alpha}\{g\} = \mathscr{F}_{\alpha+\beta}\{g\}. \tag{29}$$

上式表明广义傅里叶变换算符是可易的. 特别当 $\beta=-\alpha$ 时我们得

$$\mathscr{F}_{\alpha}\mathscr{F}_{-\alpha}\{g\} = \mathscr{F}_{-\alpha}\mathscr{F}_{\alpha}\{g\} = \mathscr{F}_{0}\{g\} = g. \tag{30}$$

这样我们进一步证明 $\mathscr{F}_{-\alpha}$ 确是 $\mathscr{F}_{\alpha}$ 的逆变换.

(7) 周期性(periodicity)

由于在广义傅里叶变换的定义中出现 $\tan\alpha$ 及 $\sin\alpha$，所以变换关于 α 具有周期性，周期为2π，这样一来，就有以下结果：

$$\mathscr{F}_{2n\pi}\{g(\xi)\} = g(x), \tag{31}$$

$$\mathscr{F}_{(2n+1)\pi}\{g(\xi)\} = g(-x), \tag{32}$$

$$\mathscr{F}_{2n\pi+\alpha}\{g\} = \mathscr{F}_{\alpha}\{g\}. \tag{33}$$

这样一来，当 $\alpha\overline{\in}(-\pi,\pi]$时的变换 $\mathscr{F}_{\alpha}$ 均可化为主值区间内的变换. 设

$$\alpha = p\pi/2, \tag{34}$$

α 阶广义傅里叶变换还可表为 $\mathscr{F}^{(p)}\{g\}$，p 的定义域为$(-2,2]$. 图5.1为 α 及 p 的周期性示意图，其中 $\alpha=\pi/2$或 $p=1$表示常规傅里叶变换，即 $\mathscr{F}_{\pi/2}$或 $\mathscr{F}^{1}$. $\mathscr{F}^{1}$常简写为 $\mathscr{F}$，与前几章的符号一致. $\alpha=-\pi/2$或 $p=-1$则表示常规的傅里叶逆变换，即 $\mathscr{F}_{-\pi/2}$或 $\mathscr{F}^{-1}$. 在(30)式中令 $\alpha=\beta=\pi/2$，即 $p_1=p_2=1$，得到

$$\mathscr{F}^{(1)}\mathscr{F}^{(1)}\{g(\xi)\} = \mathscr{F}^{(2)}\{g(\xi)\}$$

$$= \mathscr{F}_{\pi}\{g(\xi)\} = g(-x). \tag{35}$$

我们知道，$\mathscr{F}\{g\}$即常规的傅里叶变换得到的输入图像的傅里叶谱，可以用$2f$ 系统实现；而(35)式中的 $\mathscr{F}^{(2)}\{g\}$则表示两次

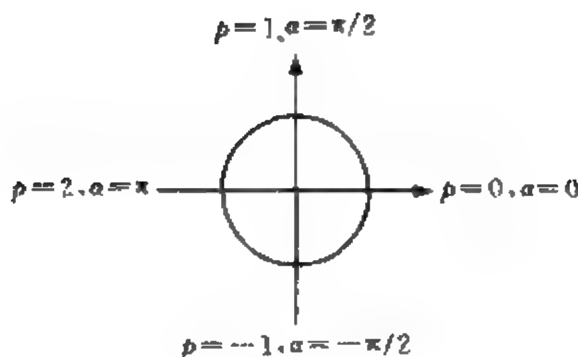


图5.1　广义傅里叶变换的周期性

傅里叶变换，得到输入图像的倒像，可以用4f系统实现，它们都是广义傅里叶变换的特例．在节5.3及节5.5中我们将研究在$\alpha\neq\pi/2$或$p\neq1$的一般情况能否用光学系统或光学器件来实现，即广义傅里叶变换的光学实现问题．

5.2.3　广义傅里叶变换群

(9)式中的变换算符$\mathscr{F}_0$具有如下性质，对于任意的$\mathscr{F}_\alpha$，有

$$\mathscr{F}_\alpha\mathscr{F}_0=\mathscr{F}_0\mathscr{F}_\alpha=\mathscr{F}_\alpha, \tag{36}$$

因此可称为单位算符或恒等元．

对于$\mathscr{F}_\alpha$，存在$\mathscr{F}_{-\alpha}$，满足

$$\mathscr{F}_\alpha\mathscr{F}_{-\alpha}=\mathscr{F}_{-\alpha}\mathscr{F}_\alpha=\mathscr{F}_0, \tag{37}$$

即$\mathscr{F}_{-\alpha}$是$\mathscr{F}_\alpha$的逆算符或逆元．对于任意的实数α,β，我们有

$$\mathscr{F}_\alpha\mathscr{F}_\beta=\mathscr{F}_\beta\mathscr{F}_\alpha=\mathscr{F}_{\alpha+\beta}, \tag{38}$$

$\mathscr{F}_{\alpha+\beta}$依然是广义傅里叶变换算符．因此变换算符对于乘法是闭合的．

此外容易证明，对于任意的实数α,β和γ，成立结合律：

$$\mathscr{F}_\alpha(\mathscr{F}_\beta\mathscr{F}_\gamma)=(\mathscr{F}_\alpha\mathscr{F}_\beta)\mathscr{F}_\gamma=\mathscr{F}_{\alpha+\beta+\gamma}, \tag{39}$$

因而所有的广义傅里叶变换算符对于(28)所定义的乘法构成群，可称为广义傅里叶变换群．

5.3 广义傅里叶变换的本征函数

广义傅里叶变换算符 $\mathscr{F}_\alpha$ 的本征函数为

$$\psi_n(x) = H_n(x)\exp(-x^2/2), \tag{1}$$

其中 $H_n(x)$ 为 n 阶厄米多项式，$\exp(-x^2/2)$ 为高斯函数，所以 $\psi_n(x)$ 常称为高斯-厄米型函数(GH 函数)．下面我们就来证明．

$$\mathscr{F}_\alpha\{\psi_n(\alpha)\} = \left\{\frac{\exp[\mathrm{i}(\alpha - \pi/2)]}{2\pi\sin\alpha}\right\}^{1/2}\exp\left(\frac{\mathrm{i}x^2}{2\tan\alpha}\right)$$
$$\times\int_{-\infty}^{\infty}\exp\left(\frac{\mathrm{i}\xi^2}{2\tan\alpha} - \frac{\mathrm{i}x\xi}{\sin\alpha}\right)\psi_n(\xi)\mathrm{d}\xi, \tag{2}$$

厄米多项式的生成函数为[19]

$$\mathrm{e}^{-t^2+2xt} = \sum_{n=0}^{\infty}\frac{t^n}{n!}H_n(x) \quad (|t| < \infty), \tag{3}$$

两边乘以 $\exp(-x^2/2)$，再进行广义傅里叶变换，得到

$$\left\{\frac{\exp[\mathrm{i}(\alpha - \pi/2)]}{2\pi\sin\alpha}\right\}^{1/2}\exp\left(\frac{\mathrm{i}x^2}{2\tan\alpha}\right)\exp(-t^2)I = \sum_{n=0}^{\infty}\frac{t^n}{n!}\mathscr{F}_\alpha\{\psi_n(\xi)\}, \tag{4}$$

上式左边的积分

$$I = \int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}\left(1 - \frac{\mathrm{i}}{\tan\alpha}\right)\xi^2 + \left(2t - \frac{\mathrm{i}x}{\sin\alpha}\right)\xi\right]\mathrm{d}\xi. \tag{5}$$

利用节5.2积分公式(11)：

$$\int_{-\infty}^{\infty}\exp(-p^2x + qx)\mathrm{d}x = \frac{\sqrt{\pi}}{p}\exp\left(\frac{q^2}{4p^2}\right),$$

其中

$$p^2 = (1 - \mathrm{i}/\tan\alpha)/2 = \mathrm{e}^{\mathrm{i}\alpha}/2\mathrm{i}\sin\alpha,$$
$$q^2/4p^2 = -x^2/2 - \mathrm{i}x^2/2\tan\alpha + 2\mathrm{i}\sin\alpha\mathrm{e}^{-\mathrm{i}\alpha}t^2 + 2xt\mathrm{e}^{-\mathrm{i}\alpha},$$
$$\sqrt{\pi}/p = (\mathrm{i}2\pi\sin\alpha\mathrm{e}^{-\mathrm{i}\alpha})^{1/2}.$$

代入(5)得到

$$I = (\mathrm{i}2\pi\sin\alpha e^{-\mathrm{i}\alpha})^{1/2}\exp\left(-\frac{\mathrm{i}x^2}{2\tan\alpha}-\frac{x^2}{2}\right)$$
$$\times \exp(\mathrm{i}2\sin\alpha e^{-\mathrm{i}\alpha}t^2 + 2xte^{-\mathrm{i}\alpha}), \tag{6}$$

以 I 代入(4)式左边,再应用厄米函数的积分公式,得到

$$\exp(-x^2/2)\exp[-(te^{-\mathrm{i}\alpha})^2 + 2x(te^{-\mathrm{i}\alpha})]$$
$$= \sum_{n=0}^{\infty}\frac{t^n}{n!}e^{-\mathrm{i}n\alpha}H_n(x)\exp(-x^2/2), \tag{7}$$

(7)式与(4)式右边应相等,亦即

$$\sum_{n=0}^{\infty}\frac{t^n}{n!}e^{-\mathrm{i}n\alpha}\psi_n(x) = \sum_{n=0}^{\infty}\frac{t^n}{n!}\mathscr{F}_\alpha\{\psi_n(\xi)\}, \tag{8}$$

从而我们有

$$\mathscr{F}_\alpha\{\psi_n(\xi)\} = e^{-\mathrm{i}n\alpha}\psi_n(x), \tag{9}$$

可见 $\psi_n=H_n(x)\exp(-x^2/2)$是广义傅里叶变换的本征函数,本征值为 $e^{-\mathrm{i}n\alpha}$.

$\{\psi_n(x)\}$构成区间$(-\infty,\infty)$内的完备正交函数组,因此任何平方可积的函数 $g(x)$都可以用它展开:

$$g(x) = \sum_{-\infty}^{\infty}a_n\psi_n(x) = \sum_{-\infty}^{\infty}a_nH_n(x)\exp(-x^2/2), \tag{10}$$

其中系数 a_n 可以用厄米函数的正交性得到:

$$a_n = \frac{1}{2^n n!\sqrt{\pi}}\int_{-\infty}^{\infty}H_n(x)\exp(-x^2/2)g(x)\mathrm{d}x \tag{11}$$

用 $\mathscr{F}_\alpha$ 作用于(10)两边,得到

$$\mathscr{F}_\alpha\{g(\xi)\} = \sum_{n=0}^{\infty}a_n\mathscr{F}_\alpha\{\psi_n(\xi)\} = \sum_{n=0}^{\infty}a_ne^{-\mathrm{i}n\alpha}\psi_n(x), \tag{12}$$

亦即

$$\mathscr{F}_\alpha\{g(\xi)\} = \sum_{n=0}^{\infty}a_ne^{-\mathrm{i}n\alpha}H_n(x)\exp(-x^2/2), \tag{13}$$

式中 a_n 由(11)式给出.

(13)式又称广义傅里叶变换的级数表达式. 在节5.6中将讨论它在渐变折射率介质光波导中的应用.

我们知道，高斯-厄米函数正是量子力学一维谐振子的本征函数. 图 5.2给出前几阶 GH 函数的图形.

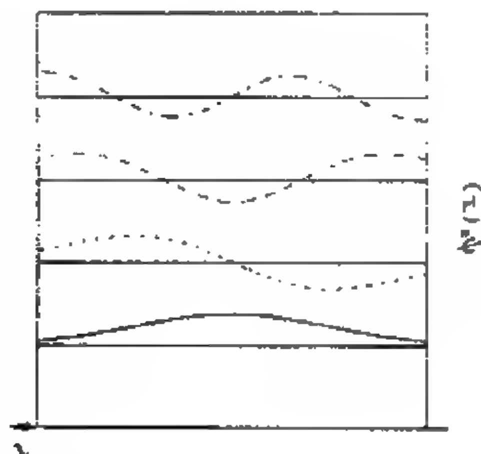


图5.2　高斯-厄米函数

5.4　用透镜系统实现广义傅里叶变换的基本光学单元

5.4.1　实现广义傅里叶变换的第一类基本光学单元

“可加性”是广义傅里叶变换的两个基本特征之一，连续执行 N 个阶数为 $\alpha_n(n=1,2,\cdots,N)$的变换的结果，相当于执行阶数为 $\alpha_1+\alpha_2+\cdots+\alpha_N$ 的一次变换，亦即

$$\mathscr{F}_\alpha = \prod_{n=1}^{N}\mathscr{F}_{\alpha_n}, \tag{1}$$

式中

$$\alpha = \sum_{n=1}^{N}\alpha_n. \tag{2}$$

举例来说，$\alpha=\pi/2$的常规傅里叶变换，既可能由一个焦距为$\tilde{f}$的透镜来实现，也可以由两个相同规格的透镜构成的透镜组来实现，它们的焦距为

$$f = \tilde{f}/\sin(\pi/4), \tag{3}$$

间距为$2d$，

$$d = f[1 - \cos(\pi/4)] = \tilde{f}\tan(\pi/8), \tag{4}$$

如图5.3所示．这一结果可以通过第一章讲述的方法，用三次菲涅耳衍射和两次透镜相位变换的方法加以证明．此外，几何光学的计算也表明，当图中 Σ_0和 Σ_2分别为系统的前、后焦面时，透镜组的焦距恰恰为$\tilde{f}$．

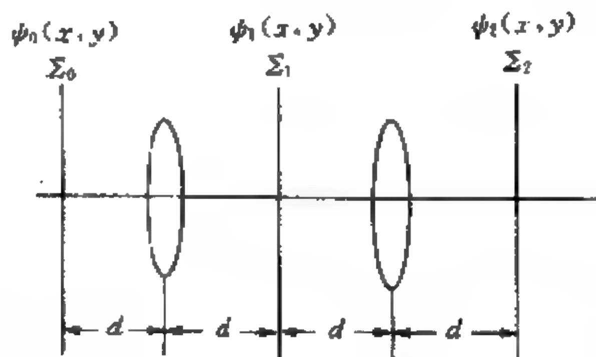


图5.3　用两个透镜实现傅里叶变换

几何光学的计算还可以证明，N 个焦距为

$$f = \tilde{f}/\sin(\pi/2N) \tag{5}$$

的透镜按图5.3的方式串联起来，间距参数

$$d = \tilde{f}[1 - \cos(\pi/2N)] = \tilde{f}\tan(\pi/4N), \tag{6}$$

则该系统的合成焦距为$\tilde{f}$，且前焦面位于第一个透镜前 d 处，后焦面位于第 N 个透镜后 d 处．但至此我们还不能证明透镜系统能实现傅里叶变换，我们必须首先证明，当单色光波通过一个透镜单元，即经过两次距离为 d 的菲涅耳衍射，并经过一次透镜相位变换，其效应相当于 $\alpha=\pi/2N$ 阶广义傅里叶变换，才能通过变换的可加性得到该系统实现傅里叶变换的普遍结论．下面我们就来研究这样一个光学单元．

广义傅里叶变换与常规的傅里叶变换的主要差别，就在于它的积分表达式（节5.2(1)式）中不仅包含线性的相位项，还包括二次相位项．该公式与第一章透镜变换的公式节1.4(18)式形式上很相似．我们重新把它记为

$$\psi(u,v)=\frac{\exp[\mathrm{i}k(d_1+d_2-f)]}{\mathrm{i}\lambda\varepsilon d_1 d_2 f}\exp\left[\frac{\mathrm{i}\pi}{\lambda\varepsilon d_1 d_2}\left(1-\frac{d_1}{f}\right)(u^2+v^2)\right]$$

$$\times\iint_{-\infty}^{\infty}\psi_0(x,y)\exp\left\{\frac{\mathrm{i}\pi}{\lambda\varepsilon d_1 d_2}\left[\left(1-\frac{d_2}{f}\right)(x^2+u^2+y^2+v^2)\right.\right.$$

$$\left.\left.-2(xu+yv)\right]\right\}\mathrm{d}x\mathrm{d}y, \tag{7}$$

式中

$$\varepsilon = 1/d_1 + 1/d_2 - 1/f, \tag{8}$$

其中的二次相位因子是由两次菲涅耳衍射及透镜相位变换的合成效果. 若

$$d_1 = d_2 = f, \tag{9}$$

则由节1.4(20)式

$$\psi(u,v)=\frac{\exp(\mathrm{i}2kf)}{\mathrm{i}\lambda f^2}\iint_{-\infty}^{\infty}\psi_0(x,y)\left[-\mathrm{i}\,\frac{2\pi}{\lambda f}(xu+yv)\right]\mathrm{d}x\mathrm{d}y$$

$$=\frac{\exp(\mathrm{i}2kf)}{\mathrm{i}\lambda f^2}\mathscr{F}\{\psi_0(\xi,\eta)\}, \tag{10}$$

即二次项消失，变成常规的傅里叶变换，这是我们已熟知的傅里叶变换的光学实现方法. 也就是说，薄透镜单元在特别的输入距离、输出距离的配置下产生了 $\alpha=\pi/2$ 的常规的傅里叶变换效应. 设想 d_1 和 d_2 不等于 f，看看在这种情况下有没有可能产生广义傅里叶变换，其阶数 $\alpha\neq\pi/2$.

将(7)式与节5.2(1)式比较，我们发现下面的条件必须成立：

$$d_1 = d_2 = d, \tag{11}$$

但 d 不一定等于 f. 再仿照(4)及(3)式，设

$$\tilde{f} = f\sin\alpha, \tag{12}$$

$$d = f(1-\cos\alpha) = \tilde{f}\tan(\alpha/2) \tag{13}$$

代入(8)式得到

$$\varepsilon d^2 = d^2(2/d - 1/f) = d(2f-d)/f = \tilde{f}\sin\alpha, \tag{14}$$

把上面的结果代入(7)式得到

$$\psi(u,v)=\frac{\exp[ik(2d-f)]}{i\lambda\tilde{f}^2}$$

$$\times\iint_{-\infty}^{\infty}\psi_0(x,y)\exp\left[\frac{i\pi(x^2+u^2+y^2+v^2)}{\lambda\tilde{f}\tan\alpha}\right.$$

$$\left.-\frac{i2\pi(xu+yv)}{\lambda\tilde{f}\sin\alpha}\right]dxdy. \tag{15}$$

引入归一化坐标(量纲为一)

$$\tilde{x}=\mu x,\quad \tilde{y}=\mu y,\quad \tilde{u}=\mu u,\quad \tilde{v}=\mu v, \tag{16}$$

其中

$$\mu=\sqrt{2\pi/\lambda\tilde{f}}=\sqrt{2\pi/\lambda f\sin\alpha}, \tag{17}$$

(15)式变成

$$\psi(\tilde{u},\tilde{v})=C\iint_{\infty}^{\infty}\psi_0(\tilde{x},\tilde{y})\exp\left[\frac{i(\tilde{x}^2+\tilde{u}^2+\tilde{y}^2+\tilde{v}^2)}{2\tan\alpha}\right.$$

$$\left.-\frac{i(\tilde{x}\tilde{u}+\tilde{y}\tilde{v})}{\sin\alpha}\right]d\tilde{x}d\tilde{y}, \tag{18}$$

式中常数

$$C=\exp[ik(2d-f)]/i\lambda\tilde{f}^2. \tag{19}$$

将(18)式与节5.2(1)式相比较,发现除积分号前的常数因子外,它就是二维 α 阶广义傅里叶变换,即

$$\psi(\tilde{u},\tilde{v})=C_0\mathscr{F}_\alpha\{\psi_0(\tilde{x},\tilde{y})\}$$

$$=C\iint_{-\infty}^{\infty}\psi_0(\tilde{x},\tilde{y})\exp\left[\frac{i(\tilde{x}^2+\tilde{u}^2+\tilde{y}^2+\tilde{v}^2)}{2\tan\alpha}\right.$$

$$\left.-\frac{i(\tilde{x}\tilde{u}+\tilde{y}\tilde{v})}{\sin\alpha}\right]d\tilde{x}d\tilde{y}. \tag{20}$$

为确保对本征函数变换时的归一化,式中 C_0和 C 有所不同.

以上讨论,证明了在条件(12)及(13)式成立时,薄透镜在单色光的照射下,可以实现二维广义傅里叶变换,它将透镜前面 d

处的输入图像 ψ_0变成它的广义傅里叶谱，形成在透镜后 d 处，如图5.4所示.

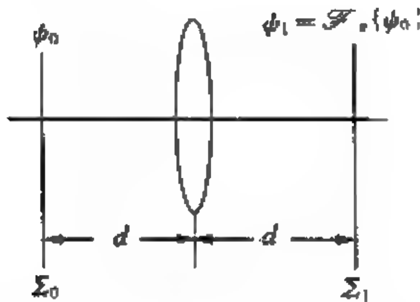


图5.4 用正透镜实现广义傅里叶变换[7]

由(12)式所定义的$\tilde{f}$称为族参数，由(13)式定义的 d 称为间距参数. 光学广义傅里叶变换表达式(15)与变换的数学定义式(节5.2(1)式)的最大差别在于光学系统中存在族参数. 很明显，只有族参数相同的光学广义傅里叶变换才能组成群，不同族参数的变换不具备可加性. 族参数仅取决于 α 及 f，然而在 α 确定后，透镜的焦距就确定了，这对光学单元按可加性组合带来许多限制，在节5.5中我们还要详细讨论.

在用透镜系统实现广义傅里叶变换时，我们宁肯不用带归一化坐标的公式(20)，而用公式(15)，即定义

$$\mathscr{F}_a\{\psi_0(x,y)\}=C_a\iint_{-\infty}^{\infty}\psi_0(x,y)\exp\left[\frac{\mathrm{i}\pi(x^2+u^2+y^2+v^2)}{\lambda\tilde{f}\tan\alpha}\right.$$
$$\left.-\frac{\mathrm{i}2\pi(xu+yv)}{\lambda\tilde{f}\sin\alpha}\right]\mathrm{d}x\mathrm{d}y, \tag{21}$$

其中积分号前的系数

$$C_a=\frac{\exp(\mathrm{i}2kd)}{\mathrm{i}\lambda\tilde{f}\sin\alpha}=\frac{\exp[\mathrm{i}2k\tilde{f}\tan(\alpha/2)]}{\mathrm{i}\lambda\tilde{f}\sin\alpha}, \tag{22}$$

与数学定义中的归一化系数并不相同，但由于我们只能探测光强的分布，因此这一差别并不带来实质性的影响.

广义傅里叶的阶数还有另一种定义，在光学中常常使用. 令

$$p=\frac{\alpha}{\pi/2}=\frac{2\alpha}{\pi}, \tag{23}$$

则(21)式还可表为

$$\mathscr{F}^{(p)}\{\psi_0(x,y)\} = C_p \iint_{-\infty}^{\infty} \psi_0(x,y) \exp\left[\frac{\mathrm{i}\pi(x^2+u^2+y^2+v^2)}{\lambda \tilde{f} \tan(p\pi/2)} - \frac{\mathrm{i}2\pi(xu+yv)}{\lambda \tilde{f} \sin(p\pi/2)}\right] \mathrm{d}x\mathrm{d}y, \tag{24}$$

(21)及(24)式都是经常运用的.

显然，$\mathscr{F}_{\pi/2}$或$\mathscr{F}^{(1)}$即常规的傅里叶变换，$\mathscr{F}_{-\pi/2}$或$\mathscr{F}^{(-1)}$即傅里叶逆变换，分别按习惯的记法，记为$\mathscr{F}$及$\mathscr{F}^{-1}$.

常规傅里叶变换仅能用正透镜实现，当我们把它推广到广义傅里叶变换时，用负透镜同样能实现广义傅里叶变换. 在(12)式中，当$f<0$时，保持$\tilde{f}>0$，则$\alpha<0$，我们得到负阶数的广义傅里叶变换，它可以用图5.5所示的负透镜单元实现. 由(13)式，当$\tilde{f}>0$，$\alpha<0$时，$d<0$，表示输入平面Σ_0在透镜右方，输出平面Σ_1在透镜左方. 当在Σ_0上输入二维图像$\psi_0(x,y)$时，它的α阶广义傅里叶变换谱出现在Σ_1处，阶数α由下式决定：

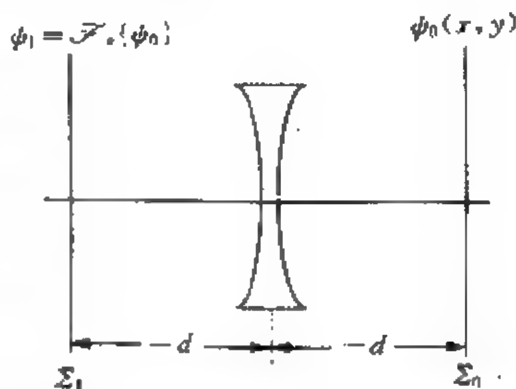


图5.5　用负透镜实现广义傅里叶变换[7]

$$\cos\alpha = 1 - d/f, \tag{25}$$

在这里d和f都是负数. 不过在这种情况下，ψ_0和ψ_1分别是光学中的虚物和虚像(参见图5.9).

5.4.2　实现广义傅里叶变换的第二类基本光学单元

图5.4或图5.5所示的光学单元通过两次菲涅耳衍射及一次透镜相位变换实现广义傅里叶变换，称为第一类基本光学单元. 可以实现广义傅里叶变换的第二类基本光学单元如图5.6所示，两个规格相同的正透镜焦距为f，间距为d，在紧贴第一个透镜前放置

输入图像ψ_0，在紧靠第二个透镜后观察输出图像ψ_1.

设在两个透镜之间光波的传播遵循菲涅耳衍射的规律，则在Σ_1面上的场

$$\psi_1(u,v)=C\iint_{-\infty}^{\infty}\psi_0(x,y)\exp\left[\frac{\mathrm{i}\pi(f-d)}{\lambda fd}(x^2+u^2+y^2+v^2)\right.$$

$$\left.-\frac{\mathrm{i}2\pi(xu+yv)}{\lambda d}\right]\mathrm{d}x\mathrm{d}y,\tag{26}$$

上式中C为常数相位因子. 按(25)及(13)式设α及d，并规定d的方向从Σ_0指向Σ_1，向右为正，则有

$$\psi_1(u,v)=C\iint_{-\infty}^{\infty}\psi_0(x,y)\exp\left[\frac{\mathrm{i}\pi\cos\alpha(x^2+u^2+y^2+v^2)}{\lambda f(1-\cos\alpha)}\right.$$

$$\left.-\frac{\mathrm{i}2\pi(xu+yv)}{\lambda f(1-\cos\alpha)}\right]\mathrm{d}x\mathrm{d}y.\tag{27}$$

将上式与(15)式相比较，发现只要设

$$d=f(1-\cos\alpha)=\tilde{f}\sin\alpha,\tag{28}$$

即

$$\tilde{f}=f\frac{1-\cos\alpha}{\sin\alpha}=f\tan\left(\frac{\alpha}{2}\right),\tag{29}$$

则有

$$\begin{aligned}\psi_1(u,v)&=C\iint_{\infty}^{\infty}\psi_0(x,y)\exp\left[\frac{\mathrm{i}\pi(x^2+u^2+y^2+v^2)}{\lambda\tilde{f}\tan\alpha}\right.\\&\qquad\left.-\frac{\mathrm{i}2\pi(xu+yv)}{\lambda\tilde{f}\sin\alpha}\right]\mathrm{d}x\mathrm{d}y,\\&=C\mathscr{F}_{\alpha}\{\psi_0\}.\end{aligned}\tag{30}$$

亦即第二类基本单元也能实现广义傅里叶变换，应注意两种基本光学单元的族参数的定义(12)及(29)式不相同.

将图5.6中两个正透镜均改为负透镜，并设f和α均为负值，此时$\tilde{f}$仍为正值，而间距

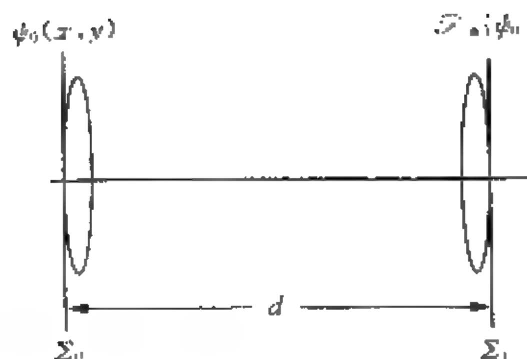


图5.6 第二类基本光学单元[7]

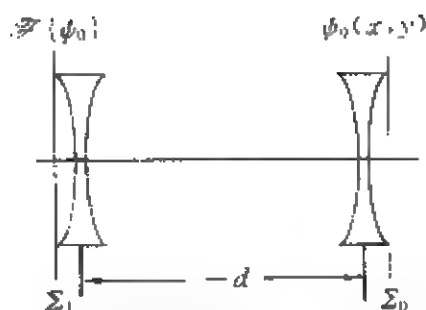


图5.7 使用负透镜的第二类基本光学单元

$$d = f(1-\cos\alpha) = \tilde{f}\sin\alpha < 0, \tag{31}$$

上式表示 ψ_0在 ψ_1左面，如图5.7所示，它能实现负阶数的广义傅里叶变换. 由此可知，用两个透镜构成的系统也能实现广义傅里叶变换.

5.5 基本光学单元的组合

在节5.4中讨论用透镜或透镜组实现广义傅里叶变换时，我们引入了族参数

$$\tilde{f} = f\sin\alpha = f\sin(p\pi/2) \quad (第一类单元), \tag{1}$$

及

$$\tilde{f} = f\frac{1-\cos\alpha}{\sin\alpha} = f\tan\left(\frac{\alpha}{2}\right) = f\tan\left(\frac{p\pi}{4}\right) \quad (第二类单元). \tag{2}$$

显然，同一类型的广义光学傅里叶算符，仅当族参数相等时才有可加性：

$$\mathscr{F}^{(p_2)}\mathscr{F}^{(p_1)} = \mathscr{F}^{(p_1+p_2)}, \tag{3}$$

换句话说，族参数相同的广义光学傅里叶算符属于同一群. (3)式暗示同属一群的光学广义傅里叶算符对应的光学单元具有互相组合成复杂系统的性能. 我们先讨论第一类光学单元的组合.

当$\tilde{f}$取某一常数时，p 或 α 可取两个值：

$$\alpha_1 = \alpha \quad (p_1 = p = 2\alpha/\pi), \tag{4}$$

$$\alpha_2 = \pi - \alpha \quad (p_2 = 2 - p), \tag{5}$$

由此导出两个不同的 d：

$$d_1 = f(1 - \cos\alpha) = f\left(1 - \cos\frac{p\pi}{2}\right) \tag{6}$$

以及

$$d_2 = f(1 + \cos\alpha) = f\left(1 + \cos\frac{p\pi}{2}\right). \tag{7}$$

下面我们举一些例子. 图5.8所示的系统由两个焦距均为 f 的透镜单元构成，它们的规格如下：

第一类光学单元：

$$d_{11} = d_{12} = d_1 = f(1 - \cos\alpha) = f\left(1 - \cos\frac{p\pi}{2}\right), \tag{8}$$

第二类光学单元：

$$d_{21} = d_{22} = d_2 = f(1 + \cos\alpha) = f\left(1 + \cos\frac{p\pi}{2}\right). \tag{9}$$

$$p_1 = p, \quad \alpha_1 = \alpha, \quad p_2 = 2 - p, \quad \alpha_2 = \pi - \alpha. \tag{10}$$

两个透镜的间距为

$$d_{12} + d_{21} = d_1 + d_2 = 2f, \tag{11}$$

它们的共同族参数为

$$\tilde{f} = f\sin\alpha_i = f\sin\left(\frac{p_i\pi}{2}\right) \quad (i = 1,2). \tag{12}$$

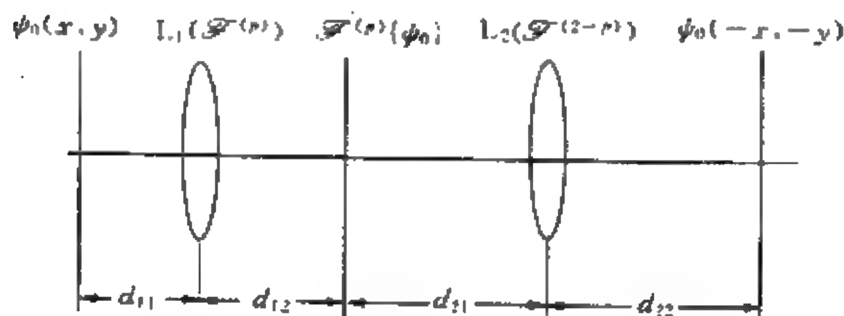


图5.8 两个相同透镜组成的广义傅里叶变换系统

设在 Σ_0平面输入图像 $\psi_0(x,y)$，则在单色光的照射下，第一类光学单元对 ψ_0进行阶数为 p_1的广义傅里叶变换，$\psi_1=\mathscr{F}^{(p_1)}\{\psi_0\}$；第二类光学单元再对 ψ_1进行阶数为 p_2的广义傅里叶变换，$\psi_2=\mathscr{F}^{(p_2)}\{\psi_1\}$，最后在输出平面 Σ_2上得到输出图像 $\psi_2(x,y)$，且有

$$\begin{aligned}\psi_2&=\mathscr{F}^{(p_2)}\mathscr{F}^{(p_1)}\{\psi_0(x,y)\}=\mathscr{F}^{(p_1+p_2)}\{\psi_0(x,y)\}\\&=\mathscr{F}^{(2)}\{\psi_0(x,y)\}=\psi_0(-x,-y),\end{aligned}\tag{13}$$

表示最终得到输入图像的倒像．当 $p=1$时我们得到4f 系统，其区别在于，在上述系统中，谱平面 Σ_1上呈现输入信号的广义傅里叶谱．而在4f 系统中，谱平面上呈现输入信号的傅里叶谱．

一个稍微复杂一点的系统包含了负透镜，见图5.9，它由三个

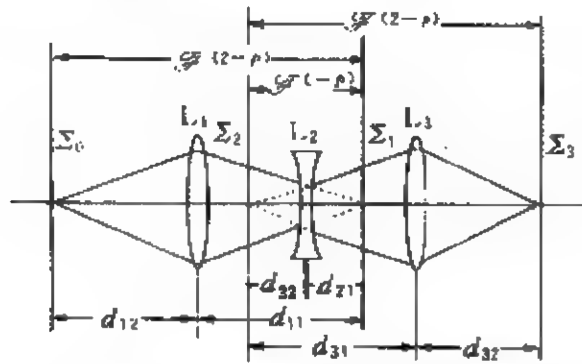


图5.9　包含负透镜的广义傅里叶变换系统[8]

第一类单元构成，L_1和 L_3为同样规格的正透镜，焦距为 f，而 L_2为负透镜，焦距为 $-f$，它们具有共同的族参数为

$$\tilde{f}=f\sin\alpha_1=f\sin\alpha_3=(-f)\sin\alpha_2.\tag{14}$$

上式成立的条件为

$$\left.\begin{aligned}&\alpha_1=\alpha_3=\pi-\alpha\quad(\alpha<\pi/2),\\&\alpha_2=-\alpha,\end{aligned}\right\}\tag{15}$$

亦即

$$\left.\begin{aligned}&p_1=p_3=2-p,\\&p_2=-p.\end{aligned}\right\}\tag{16}$$

对于中间的负透镜，我们有

$$d_{21} = d_{22} = (-f)(1 - \cos\alpha) = -f\left(1 - \cos\frac{p\pi}{2}\right), \quad (17)$$

d_{21},d_{22}为负值,表示第二类单元的输入平面位于透镜右方 Σ_1处,而输出平面位于透镜左方 Σ_2处,关于透镜对称分布.

对于两个正透镜,则有

$$d_{i1} = d_{i2} = f(1 - \cos\alpha) = f\left(1 + \cos\frac{p\pi}{2}\right) \quad (i = 1,3), \quad (18)$$

第一个光学单元的输出平面为 Σ_1,恰为第二个单元的输入平面;而第二个单元的输出平面为 Σ_2,又是第三个单元的输入平面.

由于具有共同的族参数,因此相应的广义傅里叶变换具有可加性:

$$\mathscr{F}^{(p_3)}\mathscr{F}^{(p_2)}\mathscr{F}^{(p_1)}\{\psi_0\} = \mathscr{F}^{(p_1+p_2+p_3)}\{\psi_0\} = \mathscr{F}^{(4-3p)}\{\psi_0\}, \quad (19)$$

特别当 $p=2/3$时我们有

$$\mathscr{F}^{(p_1+p_2+p_3)}\{\psi_0(x,y)\} = \mathscr{F}^{(2)}\{\psi_0(x,y)\} = \psi_0(-x,-y), \quad (20)$$

同样得到倒像.

上例表明,正、负透镜可以适当组合,其条件是它们有共同的族参数,且正、负透镜焦距的绝对值相等.

事实上,属于同一群的广义傅里叶算符只要求族参数$\tilde{f}$相同,而 f 和 α 均可不同. 例如设

$$f_1\sin\alpha_1 = f_2\sin\alpha_2 = \tilde{f}, \quad (21)$$

两个单元对应的广义傅里叶算符分别为 $\mathscr{F}_{\alpha_1}$和 $\mathscr{F}_{\alpha_2}$,它们仍属于同一群,具有可加性:

$$\mathscr{F}_{\alpha_1}\mathscr{F}_{\alpha_2} = \mathscr{F}_{\alpha_1+\alpha_2}, \quad (22)$$

所以两个光学单元之可以串接,在条件(21)满足时 α_i 有两个解,分别记为

$$\left.\begin{aligned} \alpha_1' &= \alpha_1 \quad 及 \quad \alpha_1'' = \pi - \alpha_1, \\ \alpha_2' &= \alpha_2 \quad 及 \quad \alpha_2'' = \pi - \alpha_2. \end{aligned}\right\} \quad (23)$$

与之相应，有两组可能的 d：

$$\left.\begin{aligned} d_1' &= f_1(1-\cos\alpha_1) \quad 及 \quad d_1'' = f_1(1+\cos\alpha_1), \\ d_2' &= f_2(1-\cos\alpha_2) \quad 及 \quad d_2'' = f_2(1+\cos\alpha_2). \end{aligned}\right\} \tag{24}$$

由上式还可得到

$$d_i' d_i'' = f_i^2(1-\cos^2\alpha_i) = f_i^2\sin^2\alpha_i = \tilde{f}^2 \quad (i=1,2), \tag{25}$$

亦即属于同一族的 N 个光学单元必须满足的条件为

$$d_1' d_1'' = d_2' d_2'' = \cdots = d_N' d_N'' = \tilde{f}^2. \tag{26}$$

我们举两个例子，设用两个第一类光学单元构成系统，使第一个透镜的输出平面(即系统的“谱平面”)的复振幅分布为输入图形 ψ_0的 $\mathscr{F}_{\pi/6}\{\psi_0\}$，而系统的输出平面为 ψ_0的常规傅里叶变换，即

$$\left.\begin{aligned} &\alpha_1 = \pi/6, \quad p_1 = 1/3, \\ &\tilde{f} = f_1\sin\alpha_1 = f_1/2, \\ &d_1 = f_1(1-\cos\alpha_1) = (2-\sqrt{3})f_1/2, \end{aligned}\right\} \tag{27}$$

$$\left.\begin{aligned} &\alpha_2 = \pi/2 - \alpha_1 = \pi/3, \quad p_2 = 2/3, \\ &f_2 = \tilde{f}/\sin\alpha_2 = \sqrt{3}f_1/3, \\ &d_2 = f_2(1-\cos\alpha_2) = \sqrt{3}f_1/6. \end{aligned}\right\} \tag{28}$$

构成的光学系统如图5.10所示. 几何光学的计算表明该系统的输

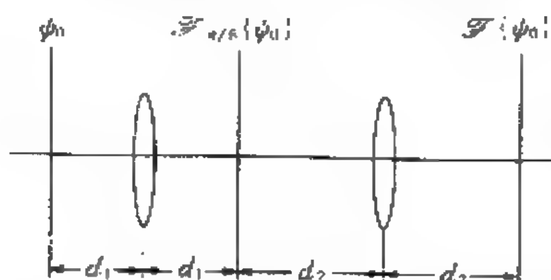


图5.10 用两个不同规格透镜的组合实现常规傅里叶变换

出平面确是系统的焦平面. 倘要求输出图像为 $\psi_0(-x,-y)$(倒像)，即要求

$$\alpha_1 + \alpha_2 = \pi, \tag{29}$$

则有

$$\sin\alpha_2 = \sin(\pi - \alpha_1) = \sin\alpha_1, \tag{30}$$

从而有

$$f_1 = f_2. \tag{31}$$

亦即两个第一类单元串联，并要求输出图像为输入图像的倒像时，两个透镜焦距相等，但 α 和 d 仍有两种不同的组合：

$$\left.\begin{aligned} \alpha_1' &= \alpha, \\ \alpha_2' &= \pi - \alpha, \end{aligned}\right\} \tag{32}$$

$$\left.\begin{aligned} \alpha_1'' &= \pi - \alpha, \\ \alpha_2'' &= \alpha, \end{aligned}\right\} \tag{33}$$

相应有

$$\left.\begin{aligned} d_1' &= f(1 - \cos\alpha), \\ d_2' &= f(1 + \cos\alpha), \end{aligned}\right\} \tag{34}$$

$$\left.\begin{aligned} d_1'' &= f(1 + \cos\alpha), \\ d_2'' &= f(1 - \cos\alpha). \end{aligned}\right\} \tag{35}$$

注意在两种情况下，透镜的间隔均为$2f$. 用几何光学容易验证在任一情形下，输出平面与输入平面共轭，且放大率为-1(倒像). 由光线的回溯性定律立即可以把这一结论推广到第二种情况.

从理论上讲，N 个相同规格的第一类单元串联起来，如每个单元对应的变换为 $\mathscr{F}^{(1/N)}$，则整个系统对应的变换为常规傅里叶变换. 由此可见，单个薄透镜并不是能实现常规傅里叶变换的唯一光学模型. 一般来讲，$\sum\alpha=\pi/2$的透镜组合均能实现常规傅里叶变换. 然而随着 N 的增大，光能损失也会变大，且杂散光的效应会变得越来越严重.

5.6 用自聚焦效应光波导实现广义傅里叶变换

5.6.1 渐变折射率介质波导的自聚焦效应

圆柱形渐变折射率波导的折射率具有下面的函数形式

$$n^2(r) = n_1^2[1 - (n_2/n_1)g(r)], \tag{1}$$

柱坐标系 $r\theta z$ 的 z 轴与波导的对称轴一致，又称光轴. $g(r)$为 r 的函数. 我们通常只考虑平方折射率介质(以下仍称为渐变折射率介质)的情形，此时有

$$\left.\begin{aligned} n^2(r) &= n_1^2[1-(n_2/n_1)r^2], \quad r<a, \\ n(r) &\approx \text{常数}, \quad r\geqslant a, \end{aligned}\right\} \tag{2}$$

其中 n_1, n_2是两个参数，a 是波导的半径. 图5.11给出 $n(r)$随 r 的变化，当 $r\geqslant a$ 时 n 趋于常数，即波导的包层. 我们仅考虑 $r\ll a$ 的近轴情况.

设单色光波沿 z 轴方向穿过相距为 Δz 的两个平面 $z=z_0$及 $z=z_0+\Delta z$，在 $z=z_0$上光波的复振幅分布为 ψ_0，则在 $z=z_0+\Delta z$ 上光波复振幅的分布为

$$\begin{aligned} \psi_1(x,y) &= \psi_0(x,y)\exp(\mathrm{i}2\pi n\Delta z/\lambda) \\ &\approx \psi_0(x,y)\exp(\mathrm{i}2\pi n_1\Delta z/\lambda_1) \\ &\quad\times \exp(-\mathrm{i}\pi n_2\Delta z(x^2+y^2)/\lambda), \end{aligned} \tag{3}$$

式中 λ 为光波在真空中的波长，并假定 Δz 很小. 将此公式与薄透镜的相位效应公式(节1.4(7)式)

$$t_l = \frac{1}{f}\exp[\mathrm{i}k(\Delta-f)]\exp\left[-\frac{\mathrm{i}k(x^2+y^2)}{2f}\right] \tag{4}$$

相比较，我们发现它们都具有关于坐标的二次相位因子，而它恰恰是球面波的近轴表达式，所以一个厚度为 Δz 的很薄的介质波导，相当于一个薄透镜，其等效焦距为$1/(n_2\Delta z)$. 这样一来，渐变介质折射率波导应具有对平行入射光的聚焦效应. 下面我们就来证明.

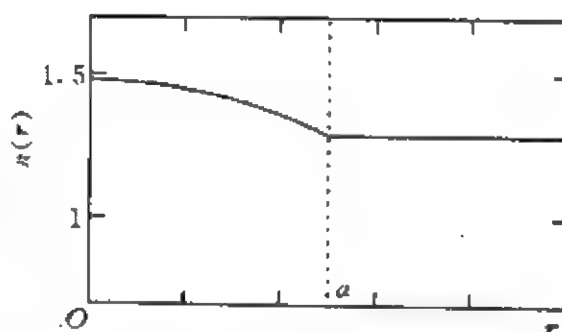


图5.11　渐变折射率波导中折射率随 r 的变化

设 A 和 B 是介质中光线经过的两点，它们之间的光程

$$\psi = \int_A^B n\mathrm{d}l = \int_{(L)} n\mathrm{d}l. \tag{5}$$

该式是沿光线传播路径 L 的积分. 根据费马原理,上式的变分为0,即

$$\delta\psi = \int_{(L)} (\delta n\mathrm{d}l + n\delta\mathrm{d}l) = 0. \tag{6}$$

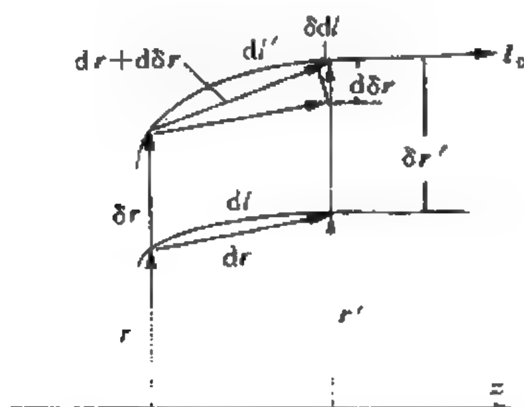


图 5.12

参见图5.12,设 $\delta\boldsymbol{r}$ 是变分时光线路径的位移,则有

$$\delta n = \delta\boldsymbol{r}\cdot\nabla n, \tag{7}$$

$$\delta\mathrm{d}l = \mathrm{d}\delta\boldsymbol{r}\cdot\boldsymbol{l}_0, \tag{8}$$

式中 $\boldsymbol{l}_0$是光线路径在给定点的单位切线向量. 代入(6)式,运用分部积分法,得到

$$\begin{aligned}\delta\psi &= \int_A^B \delta\boldsymbol{r}\cdot\nabla n\mathrm{d}l + \int_A^B n\boldsymbol{l}_0\cdot\mathrm{d}\delta\boldsymbol{r} \\ &= \int_A^B \delta\boldsymbol{r}\cdot\nabla n\mathrm{d}l - \int_A^B \delta\boldsymbol{r}\cdot\mathrm{d}(n\boldsymbol{l}_0) + n\boldsymbol{l}_0\cdot\delta\boldsymbol{r}\Big|_A^B,\end{aligned} \tag{9}$$

由于 $\delta\boldsymbol{r}$ 在两个端点 A,B 处为0,上式第三项消失. 又因为

$$\mathrm{d}\boldsymbol{r} = \boldsymbol{l}_0\mathrm{d}l, \tag{10}$$

所以有

$$\delta\psi = \int_A^B (\delta \boldsymbol{r} \cdot \nabla n)\mathrm{d}l - \int_A^B \delta \boldsymbol{r} \cdot \frac{\mathrm{d}}{\mathrm{d}l}\left(n\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}l}\right)\mathrm{d}l$$

$$= \int_A^B \delta \boldsymbol{r} \cdot \left(\nabla n - \frac{\mathrm{d}}{\mathrm{d}l} n \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}l}\right)\mathrm{d}l. \tag{11}$$

因此有

$$\frac{\mathrm{d}}{\mathrm{d}l}\left(n\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}l}\right) = \nabla n. \tag{12}$$

将 $n(r)$ 的表达式(2)代入上式右边．考虑到近轴近似，$\mathrm{d}/\mathrm{d}l \approx \mathrm{d}/\mathrm{d}z$，上式化作

$$\frac{\mathrm{d}n}{\mathrm{d}z}\frac{\mathrm{d}r}{\mathrm{d}z} + n\frac{\mathrm{d}^2 r}{\mathrm{d}z^2} = - n_2 r, \tag{13}$$

由于 n 不随 z 的变化而变化，上式左边第一项为0．用 n_1 近似代替左边第二项中的 n，得到微分方程

$$\frac{\mathrm{d}^2 r}{\mathrm{d}z^2} + \frac{n_2}{n_1} r = 0. \tag{14}$$

上式的解为

$$r(z) = r_0 \cos(\sqrt{n_2/n_1}\, z) + r_0' \sin(\sqrt{n_2/n_1}\, z), \tag{15}$$

式中 $r_0 = r(0)$，$r_0' = r'(0)$ 分别是 r 和 $\mathrm{d}r/\mathrm{d}z$ 在端口 $z=0$ 的值．设一束平行于 z 轴的光线照射 $z=0$ 端面，则 $r_0'=0$，

$$r(z) = r_0 \cos(\sqrt{n_2/n_1}\, z), \tag{16}$$

上式表明光线在渐变折射率波导中的传播路径具有周期性，如图5.13所示．设一个周期的长度为 $2L$，则有

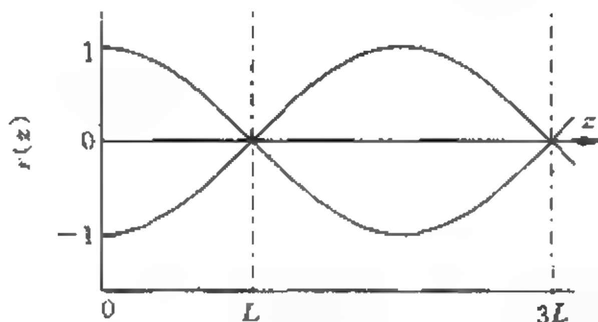


图5.13　光线在渐变折射率介质波导中的传播

$$2L = \pi\sqrt{n_1/n_2},\tag{17}$$

而焦距为

$$L = \frac{\pi}{2}\sqrt{\frac{n_1}{n_2}},\tag{18}$$

这就是自聚焦透镜的原理.

在节5.5中曾讲过，常规的傅里叶变换效应可以由 N 个第一类光学基本单元串联而成的系统来实现. 随着 N 的变大，每个单元透镜的焦距 $f=\tilde{f}/\sin(\pi/2N)$ 也随之变大，式中 $\tilde{f}$ 是系统的焦距. 因此每个薄透镜也就更接近于一个平板. 这样一系列接近密接的透镜组与上述自聚焦光波导薄层系列非常相似. 因此我们可以设想一段长度为 pL 的渐变折射率光波导能够实现 p 阶广义傅里叶变换. 下面我们将从光波在波导中传播的波动方程入手来研究这一问题.

5.6.2 光波在渐变折射率介质波导中的传播

为简单起见，我们只讨论一维情况，其结论可以直接推广到二维情况. 我们仅考虑 TE 波. 设光波沿 z 轴方向传播，电场在 y 方向振动，用 $E(x,z)$ 表示，则它应满足亥姆霍兹方程

$$\left(\frac{\partial^2}{\partial x^2} + k_x^2\right)E(x,z) = 0,\tag{19}$$

式中 k_x 是 x 方向的波矢量的值，

$$k_x^2 = k^2 - k_z^2.\tag{20}$$

令 k_0 为真空中的 k，$k_z=\beta$，称为传播常数，n 为折射率，则有

$$k^2 = n^2k_0^2,\tag{21}$$

代入(19)式，得到

$$\frac{\partial^2 E}{\partial x^2} + (n^2k_0^2 - \beta^2)E = 0.\tag{22}$$

在二次渐变折射率介质中，

$$n(x) \approx n_1 - n_2x^2/2,\tag{23}$$

这种情况相当于量子力学中无限深势阱中的一维谐振子. 将上式代入(22)式得到

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}\xi^2} + (\mu - \xi^2)\psi = 0, \tag{24}$$

式中

$$\left.\begin{aligned} &\psi = \alpha^2 E, \\ &\xi = \alpha x, \\ &\mu = (n_1^2 k_0^2 - \beta^2)/\alpha^2, \\ &\alpha = \sqrt{n_1 n_2} k_0. \end{aligned}\right\} \tag{25}$$

从(24)式解出

$$E_n(x) = N_n H_n(\sqrt{2}\, x/\omega)\exp(-x^2/\omega^2) \quad (n = 0,1,\cdots), \tag{26}$$

式中 H_n 为 n 阶厄米多项式. 由此可见,在渐变折射率波导中传播的本征模就是 GH 函数,恰恰是广义傅里叶变换的本征态.

在波导中的场为

$$E_n(x,z) = N_n H_n(\sqrt{2}\, x/\omega)\exp(-x^2/\omega^2)\exp(\mathrm{i}\beta_n z), \tag{27}$$

式中

$$N_n = n_1\left[1 - \frac{1}{n_1 k_0}\left(n + \frac{1}{2}\right)\sqrt{\frac{n_2}{n_1}}\right], \tag{28}$$

$$\beta_n = N_n k_0 = k_0\left[n_1 - \frac{1}{k_0}\left(n + \frac{1}{2}\right)\sqrt{\frac{n_2}{n_1}}\right], \tag{29}$$

ω 称光斑尺寸,

$$\omega^2 = \frac{\lambda}{\pi\sqrt{n_1 n_2}} = \frac{2}{k_0\sqrt{n_1 n_2}}, \tag{30}$$

式中 λ 为真空中的波长.

5.6.3 用渐变介质波导实现广义傅里叶变换

引入量纲一的归一化坐标

$$\tilde{x} = \sqrt{2}\, x/\omega, \quad \tilde{\xi} = \sqrt{2}\, \xi/\omega. \tag{31}$$

并记 $E_n(x)$ 为 $\psi_n(\tilde{x})$，由于 ψ_n 是$(-\infty,\infty)$区间的本征函数，所以任何平方可积的函数 $g(\tilde{x})$都可以用 $\psi_n(\tilde{x})$展开：

$$g(\tilde{x}) = \sum_{n=0}^{\infty} a_n \psi_n(\tilde{x}) = \sum_{n=0}^{\infty} a_n H_n(\sqrt{2}\, x/\omega) \exp(-x^2/\omega^2), \tag{32}$$

系数

$$a_n = \frac{1}{h_n} \int_{-\infty}^{\infty} g(\tilde{x}) \psi_n(\tilde{x}) \mathrm{d}\tilde{x}, \tag{33}$$

$$h_n = 2^n n! \sqrt{\pi}. \tag{34}$$

设我们在 $z=0$处放置由 $g(\tilde{x})=g(\sqrt{2}/\omega x)$表达的一维图像，则在 $z=pL$ 处的场为

$$\begin{aligned} g(\tilde{x}, pL) &= \sum_{n=0}^{\infty} a_n \psi_n(\tilde{x}) \exp(\mathrm{i}\beta_n pL) \\ &= \sum_{n=0}^{\infty} a_n H_n\left(\frac{\sqrt{2}}{\omega} x\right) \exp\left(-\frac{x^2}{\omega^2}\right) \exp(\mathrm{i}\beta_n pL). \end{aligned} \tag{35}$$

上面曾得到结论：长度为 $L(p=1)$的光波导能实现光波的聚焦. 我们将讨论在 $p\neq 1$时，长度为 pL 的光波导对光波的变换关系.

注意到焦距 L 的表达式(18)，就有

$$\begin{aligned} \beta_n pL &= \left[k_0 n_1 L - \left(n + \frac{1}{2}\right)\frac{\pi}{2}\right] p \\ &= \left[k_0 n_1 L - \left(n + \frac{1}{2}\right)\frac{\pi}{2}\right]\left(\frac{2\alpha}{\pi}\right) \\ &= \left[k_0 n_1 \left(\frac{2\alpha}{\pi}\right) L - \frac{\alpha}{2}\right] - n\alpha, \end{aligned} \tag{36}$$

代入(35)式，得到

$$g(\tilde{x}, pL) = A \sum_{n=0}^{\infty} a_n \mathrm{e}^{-\mathrm{i}n\alpha} \psi_n(\tilde{x}), \tag{37}$$

式中

$$A = \exp\left\{\mathrm{i}\left[k_0 n_1\left(\frac{2\alpha}{\pi}\right)L - \frac{\alpha}{2}\right]\right\}. \tag{38}$$

将广义傅里叶变换对本征函数的变换关系式(节5.2(9)式)代入(37),得到

$$\begin{aligned} g(\tilde{x}, pL) &= A\sum_{n=0}^{\infty} a_n \mathscr{F}_{\alpha}\{\psi_n(\tilde{\xi})\} \\ &= A\mathscr{F}_{\alpha}\sum_{n=0}^{\infty} a_n \psi_n(\tilde{\xi}) = A\mathscr{F}_{\alpha}\{g(\tilde{\xi})\} \\ &= A\mathscr{F}_{\alpha}\{g(\tilde{\xi}, 0)\} = A\mathscr{F}^{(p)}\{g(\tilde{\xi}, 0)\}. \end{aligned} \tag{39}$$

(39)式表明:一段长度为 pL 的渐变折射率介质波异,也能实现广义傅里叶变换. 在 $z=0$处放置图像 g,并用单色光波照射,在 $z=pL$ 处即可得到它的广义傅里叶变换谱 $\mathscr{F}^{(p)}\{g\}$. 显然,长度为 p_1L 和 p_2L 的两段光波导串接起来,必然能实现 p_1+p_2阶广义傅里叶变换,而变换算符的可加性及可易性是由光波导长度的可加性及可易性自然满足的.

这样一来,我们又找到一种实现广义傅里叶变换的光学器件. 串联的透镜组只能得到 P/Q 阶广义傅里叶变换,P 和 Q 都是正整数,而由于技术上的原因,Q 无法做到太大,所以我们只能得到阶数的步进的变化. 由于光波导中阶数正比于波导的长度,因此能够实现阶数的连续的变化.

光波导的缺点,是空间带宽积比较小,也就是说它能处理的信息量比较小.

5.7 维格纳变换

5.7.1 维格纳变换的定义

函数 $g(x)$的维格纳(Wigner)变换定义为

$$W(x,\nu) = \int_{-\infty}^{\infty} g\left(x+\frac{x'}{2}\right)g^*\left(x-\frac{x'}{2}\right)\exp(-\mathrm{i}2\pi\nu x')\mathrm{d}x'. \tag{1}$$

如果把 x,ν 分别理解为信号 g 的空间变量和空间频率变量,则维格纳变换就是信号函数 g 的空间和频谱特征的综合表现或同时反映,因此在光学信号及其他信号的处理中有广泛的应用. 维格纳变换的缺点在于它并不是线性变换,所以将它应用到线性系统中去有一定的困难.

设 $G(\nu)$ 是 $g(x)$ 的傅里叶变换,则有

$$g\left(x+\frac{x'}{2}\right)=\int_{-\infty}^{\infty}G(\nu)\exp\left[\mathrm{i}2\pi\nu\left(x+\frac{x'}{2}\right)\right]\mathrm{d}\nu, \tag{2}$$

$$g^*\left(x-\frac{x'}{2}\right)=\int_{-\infty}^{\infty}G^*(\nu)\exp\left[-\mathrm{i}2\pi\nu\left(x-\frac{x'}{2}\right)\right]\mathrm{d}\nu, \tag{3}$$

以(2)及(3)式代入(1)式得到

$$W=\iint_{-\infty}^{\infty}G(\nu')G^*(\nu'')\exp[\mathrm{i}2\pi(\nu'-\nu'')x]I\mathrm{d}\nu'\mathrm{d}\nu'', \tag{4}$$

其中的积分

$$\begin{aligned}I&=\int_{-\infty}^{\infty}\exp\left[\mathrm{i}2\pi\left(\frac{\nu'+\nu''}{2}-\nu\right)x'\right]\mathrm{d}x'\\&=2\delta(\nu''-(2\nu-\nu')),\end{aligned} \tag{5}$$

代入(4)式,经整理后得到维格纳变换在频域中的表达式

$$W(x,\nu)=\int_{-\infty}^{\infty}G\left(\nu+\frac{\nu'}{2}\right)G^*\left(\nu-\frac{\nu'}{2}\right)\exp(\mathrm{i}2\pi\nu' x)\mathrm{d}\nu'. \tag{6}$$

容易证明逆变换表达式

$$\int_{-\infty}^{\infty}W(x,\nu)\exp(\mathrm{i}4\pi\nu x)\mathrm{d}\nu=g(2x)g^*(0), \tag{7}$$

$$\int_{-\infty}^{\infty}W(x,\nu)\exp(-\mathrm{i}4\pi\nu x)\mathrm{d}x=G(2\nu)G^*(0). \tag{8}$$

维格纳变换对于 ν 和 x 的积分分别得到信号在空域和频域中的功率密度

$$\int_{-\infty}^{\infty}W(x,\nu)\mathrm{d}\nu=|g(x)|^2, \tag{9}$$

$$\int_{-\infty}^{\infty}W(x,\nu)\mathrm{d}x=|G(\nu)|^2, \tag{10}$$

因而信号的总功率 E_0 可以由下面的积分来表达：

$$\iint_{-\infty}^{\infty} W(x,\nu)\mathrm{d}x\mathrm{d}\nu = E_0. \tag{11}$$

下面让我们来讨论当信号变化时其维格纳谱的相应变化. 首先把(1)式改写成更加容易处理的形式：

$$W(x,\xi) = \int_{-\infty}^{\infty} g\left(x + \frac{x'}{2}\right) g^*\left(x - \frac{x'}{2}\right) \exp\left[-\frac{\mathrm{i}2\pi\xi x'}{\lambda \tilde{f}}\right] \mathrm{d}x', \tag{12}$$

$$W(x,\xi) = \int_{-\infty}^{\infty} G\left(\xi + \frac{\xi'}{2}\right) G^*\left(\xi - \frac{\xi'}{2}\right) \exp(\mathrm{i}2\pi\lambda \tilde{f}\xi' x) \mathrm{d}\xi', \tag{13}$$

式中 λ 为光波的波长；$\tilde{f}$ 是维格纳变换的一个特征量，以后我们会看到它相当于广义傅里叶变换中的 $\tilde{f}$；ξ 和 x 都具有长度量纲，分别是频域和空域中的长度坐标；$\xi/\lambda\tilde{f}$ 则是空间频率变量.

5.7.2 基本性质和运算法则

(1) 反演(inversion)

当

$$g(x) \longrightarrow g(-x) \tag{14}$$

时，相应的维格纳变换成为

$$\int_{-\infty}^{\infty} g\left(-x - \frac{x''}{2}\right) g^*\left(-x + \frac{x''}{2}\right) \exp\left(-\frac{\mathrm{i}2\pi\xi x''}{\lambda\tilde{f}}\right) \mathrm{d}x''. \tag{15}$$

令 $x'' = -x'$，上式变为

$$\int_{-\infty}^{\infty} g\left(-x + \frac{x'}{2}\right) g^*\left(-x - \frac{x'}{2}\right) \exp\left[-\frac{\mathrm{i}2\pi(-\xi)x'}{\lambda\tilde{f}}\right] \mathrm{d}x' = W(-x, -\xi). \tag{16}$$

可见信号在 x 空间的反演引起维格纳变换在 (x,ξ) 空间的反演(参见图5.14)：

$$W(A) \longrightarrow W(A'), \tag{17}$$

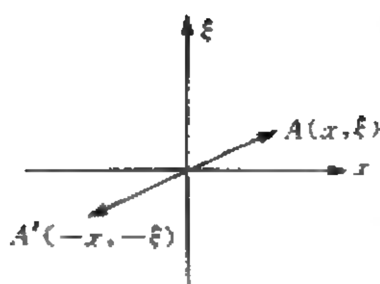


图 5.14

上述变换可以用矩阵来表示，设(x,ξ)和(x',ξ')分别是变化前后的维格纳空间坐标，

$$\begin{pmatrix} x' \\ \xi' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ \xi \end{pmatrix}, \qquad (18)$$

则反演的结果可用下式表示：

$$\begin{pmatrix} x' \\ \xi' \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ \xi \end{pmatrix} = \begin{pmatrix} -x \\ -\xi \end{pmatrix}. \qquad (19)$$

(2) 缩放(scaling)

设

$$g(x) \longrightarrow g(x/M), \qquad (20)$$

式中 M 为缩放倍率，相应的维格纳谱的变化为

$$W(x,\xi) \longrightarrow W(x/M, M\xi), \qquad (21)$$

或

$$\begin{pmatrix} x' \\ \xi' \end{pmatrix} = \begin{pmatrix} 1/M & 0 \\ 0 & M \end{pmatrix} \begin{pmatrix} x \\ \xi \end{pmatrix} = \begin{pmatrix} x/M \\ M\xi \end{pmatrix}, \qquad (22)$$

即信号坐标 x 的放大将引起维格纳空间中 x 坐标按同样倍率的放大，同时 ξ 坐标按同样比例缩小.

(3) 薄透镜的相位变换

$$g(x) \longrightarrow g(x)\exp(-\mathrm{i}\pi x^2 Q/\lambda\tilde{f}), \qquad (23)$$

在此我们设透镜的焦距

$$f = \tilde{f}/Q. \qquad (24)$$

相应的维格纳变换为

$$\int_{-\infty}^{\infty} g\left(x+\frac{x'}{2}\right) g^*\left(x-\frac{x'}{2}\right) \exp\left\{-\mathrm{i}\frac{\pi Q}{\lambda\tilde{f}}\left[\left(x+\frac{x'}{2}\right)^2 - \left(x-\frac{x'}{2}\right)^2\right]\right\} \exp\left[-\mathrm{i}\frac{2\pi\xi x'}{\lambda\tilde{f}}\right] \mathrm{d}x'$$

$$= \int_{-\infty}^{\infty} g\left(x + \frac{x'}{2}\right) g^*\left(x - \frac{x'}{2}\right) \exp\left[-\frac{i2\pi(Qx + \xi)x'}{\lambda \tilde{f}}\right] dx'$$

$$= W(x, \xi + Qx), \tag{25}$$

变换矩阵表达式则为

$$\begin{pmatrix} x' \\ \xi' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ Q & 1 \end{pmatrix} \begin{pmatrix} x \\ \xi \end{pmatrix} = \begin{pmatrix} x \\ \xi + Qx \end{pmatrix}, \tag{26}$$

在维格纳空间中表现为坐标沿平行于 ξ 轴的直线移动 Qx，见图5.15.

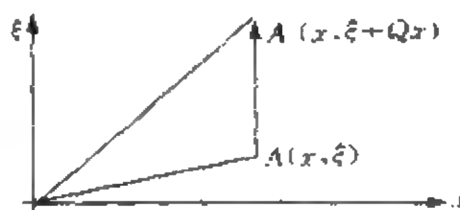


图 5.15

(4) 信号在自由空间中的传播——菲涅耳衍射

我们在频域中进行变换. 由节1.3(12)式，信号在自由空间中传播的距离为 z 时，相应傅里叶谱 G 如下变化

$$G(\xi) \longrightarrow G(\xi)\exp(-i\pi\lambda z\xi^2), \tag{27}$$

这样一来，我们就有

$$G\left(\xi + \frac{\xi'}{2}\right) \longrightarrow G\left(\xi + \frac{\xi'}{2}\right)\exp\left[-i\pi\lambda z\left(\xi + \frac{\xi'}{2}\right)^2\right], \tag{28}$$

$$G^*\left(\xi - \frac{\xi'}{2}\right) \longrightarrow G^*\left(\xi - \frac{\xi'}{2}\right)\exp\left[i\pi\lambda z\left(\xi - \frac{\xi'}{2}\right)^2\right], \tag{29}$$

代入(13)式，得到

$$W' = \int_{-\infty}^{\infty} G\left(\xi + \frac{\xi'}{2}\right) G^*\left(\xi - \frac{\xi'}{2}\right) \exp\left[i2\pi\lambda \tilde{f}\left(x - \frac{z}{\tilde{f}}\xi\right)\xi'\right] d\xi'$$

$$= W(x - R\xi, \xi), \tag{30}$$

式中

$$R = z/\tilde{f}, \tag{31}$$

相应的矩阵变换式为

$$\begin{pmatrix} x' \\ \xi' \end{pmatrix} = \begin{pmatrix} 1 & -R \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ \xi \end{pmatrix} = \begin{pmatrix} x - R\xi \\ \xi \end{pmatrix}, \tag{32}$$

在维格纳空间中相当于坐标沿平行于 x 轴的平移，见图5.16.

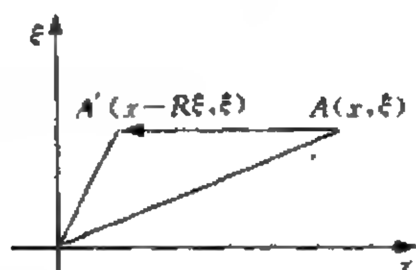


图 5.16

(5) 夫琅和费衍射

$$g(x) \longrightarrow \int_{-\infty}^{\infty} g(x_0) \exp\left(-\mathrm{i}\frac{2\pi x_0 x}{\lambda z}\right) \mathrm{d}x_0. \tag{33}$$

将它代入(12)式，得到

$$W' = \int_{-\infty}^{\infty} g\left(-\xi + \frac{x_0'}{2}\right) g^* \left[(1 - 2R)\xi - \frac{x_0'}{2}\right]$$

$$\times \exp\left\{-\mathrm{i}\frac{2\pi x}{\lambda z}[2(R-1)\xi + x_0]\right\} \mathrm{d}x_0', \tag{34}$$

式中 R 由(31)式表示. 令 $R=1$得到

$$W' = \int_{-\infty}^{\infty} g\left(-\xi + \frac{x_0'}{2}\right) g^* \left(-\xi - \frac{x_0'}{2}\right) \exp\left(-\mathrm{i}\frac{2\pi x x_0'}{\lambda \tilde{f}}\right) \mathrm{d}x_0'$$

$$= W(-\xi, x), \tag{35}$$

即

$$\begin{pmatrix} x' \\ \xi' \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ \xi \end{pmatrix} = \begin{pmatrix} -\xi \\ x \end{pmatrix}. \tag{36}$$

5.7.3 广义傅里叶变换的维格纳变换定义

在维格纳表象中，由(36)式表示的常规傅里叶变换表示绕原点旋转 $\pi/2$，见图5.17. 我们自然要问，当 $\alpha \neq \pi/2$时的旋转能否表示广义傅里叶变换？下面我们用两种方法来加以证明.

在广义傅里叶变换下，

$$g(x)\longrightarrow\int_{-\infty}^{\infty}g(x_0)\exp\left[i\pi\frac{x^2+x_0^2}{\lambda\tilde{f}\tan\alpha}-i2\pi\frac{xx_0}{\lambda\tilde{f}\tan\alpha}\right]dx_0.\quad(37)$$

$$g\left(x+\frac{x'}{2}\right)\longrightarrow\int_{-\infty}^{\infty}g(x_0)\exp\left[i\pi\frac{(x+x'/2)^2+x_0^2}{\lambda\tilde{f}\tan\alpha}\right.$$

$$\left.-i2\pi\frac{(x+x'/2)x_0}{\lambda\tilde{f}\tan\alpha}\right]dx_0.\quad(38)$$

$$g^*\left(x-\frac{x'}{2}\right)\longrightarrow\int_{-\infty}^{\infty}g^*(x_0)\exp\left[-i\pi\frac{(x-x'/2)^2+x_0^2}{\lambda\tilde{f}\tan\alpha}\right.$$

$$\left.+i2\pi\frac{(x-x'/2)x_0}{\lambda\tilde{f}\sin\alpha}\right]dx_c,\quad(39)$$

代入(12)式得到

$$W'=\iint_{-\infty}^{\infty}g(x_0')g^*(x_0'')\exp\left[i\pi\frac{x_0'-x_0''}{\lambda\tilde{f}\tan\alpha}\right.$$

$$\left.-i2\pi\frac{x(x_0'-x_0'')}{\lambda\tilde{f}\sin\alpha}\right]Idx_0'dx_0'',\quad(40)$$

式中积分的结果为 δ 函数：

$$I=2\lambda\tilde{f}\sin\delta[x_0''-(2x\cos\alpha-2\xi\sin\alpha-x_0')],\quad(41)$$

代入(40)式，经过整理，得到

$$W'=C\int_{-\infty}^{\infty}g\left(x\cos\alpha-\xi\sin\alpha+\frac{x'}{2}\right)g^*\left(x\cos\alpha\right.$$

$$\left.-\xi\sin\alpha-\frac{x'}{2}\right)\exp\left[-\frac{i2\pi}{\lambda\tilde{f}}(x\sin\alpha+\xi\cos\alpha)x'\right]dx'$$

$$=CW(x\cos\alpha-\xi\sin\alpha,\ x\sin\alpha+\xi\cos\alpha),\quad(42)$$

式中 C 为常数. 相应的矩阵变换式为

$$\begin{pmatrix}x'\\ \xi'\end{pmatrix}=\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}x\\ \xi\end{pmatrix}=\begin{pmatrix}x\cos\alpha-\xi\sin\alpha\\ x\sin\alpha+\xi\cos\alpha\end{pmatrix},\quad(43)$$

上式表明在维格纳空间中，广义傅里叶变换相当于坐标系旋转 α 角. 从另一角度来看，由于 α 阶广义傅里叶变换可以由第一类光学系统实现，它是两次菲涅耳衍射，以及在其间的透镜相位变换的结果. 设两次光波传播对应的距离为 z，透镜的焦距为 f，由(24)及(31)式

$$f=\tilde{f}/Q,\quad z=R\tilde{f}.$$

按广义傅里叶变换的参数定义，令

$$R=\tan(\alpha/2)=(1-\cos\alpha)/\sin\alpha, \tag{44}$$

$$Q=\sin\alpha, \tag{45}$$

则信号的广义傅里叶变换对应的维格纳变换可以由矩阵变换式得到

$$\begin{aligned}\begin{bmatrix}x'\\ \xi'\end{bmatrix}&=\begin{bmatrix}1 & -R\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ Q & 1\end{bmatrix}\begin{bmatrix}1 & -R\\ 0 & 1\end{bmatrix}\\ &=\begin{bmatrix}1-RQ & R(QR-2)\\ Q & 1-RQ\end{bmatrix}\\ &=\begin{bmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{bmatrix}.\end{aligned} \tag{46}$$

我们用矩阵变换式完成了第二种证明.

(42)式是用维格纳变换表示的广义傅里叶变换的定义. 它显然符合广义傅里叶变换的两个基本假设：首先，当 $\alpha=\pi/2$时我们得到常规的傅里叶变换；其次，依次施行阶数为 α_1及 α_2的两次广义傅里叶变换，在维格纳空间表现为坐标系依次旋转 α_1和 α_2，其结果是旋转了 $\alpha_1+\alpha_2$角，表明变换具有可加性.

在维格纳表象中，广义傅里叶变换($A\to A'$)及常规傅里叶变换($A\to A''$)见图5.17所示.

5.7.4 维格纳变换在光学系统设计中的应用

维格纳变换是光学系统设计的有力工具，特别是因为信号(即

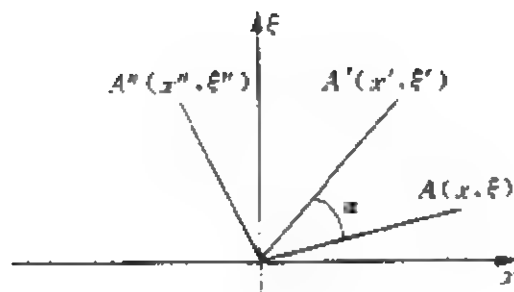


图 5.17

光学图像)的变化可用矩阵表示或用维格纳表象中的简单变换表示,因而可以用维格纳变换进行光学系统的总体设计.

例如,要设计一个焦距可以调节的光学系统,亦即要求(26)式的变换矩阵的参数 Q 可变.透镜的焦距是固定的,因而这一要求只能通过透镜间的距离的调节来达到.

我们不研究普遍的设计方法,只讨论一个具体的方案:用三个变换矩阵来达到设计指标,前后矩阵表示常规傅里叶变换,中间的矩阵是光波在自由空间中的传播,传播的距离为 z. 则有

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -R \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ -R & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ R & 1 \end{pmatrix}. \tag{47}$$

式中最后两个矩阵的积表示生成倒像,而矩阵

$$\begin{pmatrix} 1 & 0 \\ R & 1 \end{pmatrix}$$

则表示三个矩阵的积相当于一个透镜,它的合成焦距

$$f = \tilde{f}/R = (\tilde{f})^2/z \tag{48}$$

可以通过自由传播的距离 z 来调节. 在这个例子中,参数 $\tilde{f}$ 相当于单个透镜的焦距,即

$$\tilde{f} = f_1 = f_2. \tag{49}$$

该系统可以由第一类单元或第二类单元来实现,见图5.18.

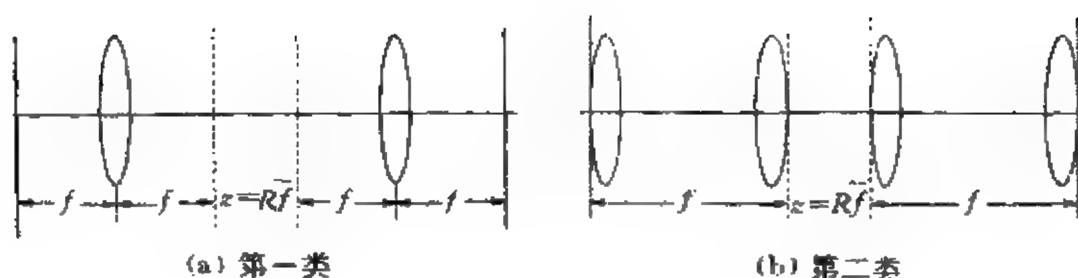


图5.18 用第一类或第二类单元构造ZOOM系统

参考文献

[1] E. U. Condon, Immersion of the Fourier transform in a continuous group of functional trnsformation, *Acad. Sci.* USA, **23**(1937), 158-164.

[2] V. Bargmann, On a Hilbert space of analytic functions and an associated integral transforms, Part I, *Commun. Pure Appl. Math.*, **14**(1961), 187-214.

[3] V. Namias, The fractional order Fourier transform and its application to quantum mechanics, *J. Inst. Maths Applics*, **25**(1980), 241-265.

[4] A. C. McBride and F. H. Kerr, On Namias fractional fourier transforms, *IMA J. Appl. Maths*, **39**(1987), 159-175.

[5] H. M. Ozaktas and D. Mendlovic, Fourier transforms of fractional order and their optical interpretation, *Opt. Commun.*, **101**(1993), 163-169.

[6] D. Mendlovic and H. M. Ozaktas, Fractional Fourier transforms and their optical implementation, *I, J. Opt. Soc. Am.*, **A 10**(1993), 1875-1881.

[7] A. W. Lohmann, Image rotationm Wigner rotation, and fractional Fourier transform, *J. Opt. Soc. Am.*, **A 10**(1993), 2181-2186.

[8] L. Bernardo and O. D. D. Soares, Fractional Fourier transform and optical sestyms, *Opt. Commun.*, **110**(1994), 517-522.

[9] A. W. Lohmann, A fake zoom lens for Fractional Fourier experiments, *Opt. Commun.*, **115**(1995), 437-443.

[10] D. Mendlovic, H. M. Ozaktas and A. W. Lohmmann, Graded-index fiber, Wigner-distribution functions, and the fractional Fourier transform, *Appl. Opt.*, **33** (1994), 6188-6193.

[11] D. Mendlovic, Z. Zalevsky, R. G. Dorsch and Y. Bitran, New signal representation based on the fractional Fourierm; definitions, J. *Opt. Soc. Am.*, **A12** (1995), 2424-2431.

[12] D. Mendlovic, Y. Bitran, R. G. Dorsch, and A. W. Lohmann, Optical fractional correlation: experimental results, J. *Opt. Soc. Am.*, **A12**(1995), 1665-1670.

[13] R. D. Dorsch, Fractional Fourier transformer of variable order based on a modular lens sestym, *Appl. Opt.*, **34**(1995), 6016-6020.

[14] L. M. Bernardo and O. D. D. Soares, Fractional Fourier transform and imageing, *J. Opt. Soc. Am.*, **A 11**(1994), 2622-2626.

[15] H. M. Ozaktas, B. Barshan, D. Mendlovic and L. Onural, Convolution, filtering, and multiplexing in fractional Fourier domain and their relation to chirp and wavelet transforms, *J. Opt. Soc. Am.*, **A11**(1994), 547-559.

[16] D. Mendlovic, Z. Zaleysky, N. Konforti and R. G. Dorsch, Incoherentfractional Fourier transform and optical implementation, *Appl. Opt.*, **34**(1995), 7615-7620.

[17] D. Mendlovic, R. G. Dorsch, A. W. Lohmann and Z. Zaleysky, Optical illustration of varied fractional Fourier-transform order and the Radon-Wigner display, *Appl. Opt.*, **35**(1996), 3925-3929.

[18] P. Pellat-Finet, Fresnel diffraction and the fractional-order Fourier transform, *Opt. Lett.*, **19**(1994), 1388-1390.

[19] 王竹溪,郭敦仁,特殊函数概论,科学出版社,1965.

第六章　光学小波变换

6.1　引　　言

傅里叶变换已成为信息处理中一个极为重要的工具，在科学和技术的许多领域中获得广泛的应用．信号 $g(x)$ 的傅里叶变换定义为

$$G(\nu)=\int_{-\infty}^{\infty} g(x)\exp(-\mathrm{i}2\pi\nu x)\mathrm{d}x. \tag{1}$$

其逆变换

$$g(x)=\int_{-\infty}^{\infty} G(\nu)\exp(\mathrm{i}2\pi\nu x)\mathrm{d}\nu, \tag{2}$$

(1)式表示信号 $g(x)$ 中频率为 ν 的成分含量为 $G(\nu)$，x 可以是时间变量或空间变量，$G(\nu)$ 则分别表示时间频率或空间频率的成分含量．(1)式表示，$G(\nu)$ 是 $g(x)$ 与表示简谐振动或平面波的指数函数 $\exp(-\mathrm{i}2\pi\nu x)$ 的积在 $-\infty$ 到 ∞ 的积分；(2)式则表示信号 $g(x)$ 可以分解为一系列的简谐振动或平面波的线性叠加，$G(\nu)$ 正是叠加运算中的权函数．

如果 $g(x)$ 是一个时域或空域中分布在 $(-\infty,\ \infty)$ 中的稳恒过程或稳定分布，则傅里叶分析给出了近乎完美的结果．然而，在自然界和科学技术中还有大量信号，它们具有局部的或定域的特性．例如语言信号、声纳信号、各种电脉冲等，这些信号只出现在一个短暂的时间间隔内，此后很快衰减到零，称快速过程或暂态过程．

图 6.1 给出这样一个信号 $S(t)$，它可以是声纳的波包，也可以是地震波．该信号的特点是在某一时刻 t 突然出现(此前可能完全没有信号)，且很快衰减到零，所以是一个很短暂的信号，可以

称为“小波”信号. 许多光学信号具有同样的特征，例如远处空中的目标、显微镜下的小物体、被鉴别的指纹等，它们不显著为零的分量只分布在有限的区域内. 上述信号称局部信号或暂态过程. 对于局部信号或暂态过程，傅里叶分析就不完全适用. 首先，我们仅对 Δt 内的时间信号感兴趣，没有必要在过去、现在及未来的无限长时间范围内对信号进行分析；类似地，在处理定域于 Δx 内的空间图像时，也没有必要对全平面内的信号进行全面的分析.

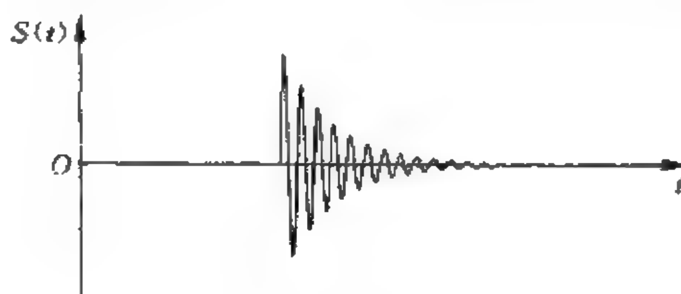


图 6.1 “小波”信号

在许多情况下，在 Δt 或 Δx 以外的信号是未知的，它可能是零，也可能是背景噪声；对它们我们不太了解，测不准，或不感兴趣. 如不加选择地把 $(-\infty,\infty)$ 内全部信号进行傅里叶处理，还可能产生较大的误差甚至错误. 此外，一个局部的信号在 Δt 或 Δx 以外较远处几乎完全等于零. 当用它们的频谱来恢复或重构这些信号时，在 Δt 或 Δx 外很远处也会出现一些非零的分量，它们一般不是信号，而是在傅里叶逆变换中频域综合不够充分而产生的噪声.

在一些课题中，我们往往不满足于了解信号在全部区间内的综合的频谱分布，而希望了解某一区间或某些区间内信号对应的频谱. 例如在地震勘探中，为了分辨分层的地层和矿床结构，我们需要在时域和频域中仔细分析不同时刻的信号在不同频谱区间中的行为，而傅里叶分析只能提供在长时间内的信号整体的频谱，显然不能满足我们的要求.

近年来发展起来的小波分析，正好克服了傅里叶分析的上述缺点. 它和傅里叶分析的一个重要区别，在于它恰恰适用于处理局部或暂态信号. 因此，小波分析成为信号分析、图像处理、数据压缩、语音信号分析等领域中的重要工具；在地震勘探信号处理、边缘探测、语音信号合成中则有特殊的用途. 近年来，小波变换与分形之间的联系引起了人们的极大兴趣[12]，小波分析已发展成为数学中的一个新兴的学科分支.

在本章中，作者从短时傅里叶变换(Short-time Fourier transform)的分析入手，引入小波变换的基本观念，然后讨论小波变换的基本性质，并介绍实现小波变换的光学系统. 对于小波变换的理论有兴趣的读者，可参阅 Chui 的经典著作[1]，以及 Li[2]等和 Sheng 等[4]撰写的综述.

6.2 从短时傅里叶变换到小波变换

6.2.1 短时傅里叶变换(STFT)[10]

为了有效地提取一个局部信号 $g(x)$的信息，我们必须引入一个局部化的变换. 所谓局部化，包含两个要素：第一，被分析的区间要有一定的宽度 Δx，我们仅对 Δx 及其附近的信息进行处理；第二，被分析的区间有一个中心坐标 x_c，当 x_c 改变时，就可以提取不同的信息.

为了实现局部化，一个有效的方案是在傅里叶变换中加一个窗函数 $w(x)$：

$$G_w(\nu,x_0)=\int_{-\infty}^{\infty}g(x)\ \exp(-i2\pi\nu x)\ w(x-x_0)\mathrm{d}x. \qquad (1)$$

在频域中，上式可表为

$$G_w(\nu,x_0)=[W(\nu)\exp(-i2\pi\nu x_0)]*G(\nu), \qquad (2)$$

式中 W 和 G 分别是 w 和 g 的傅里叶变换.

只要 $w(x)$和 $W(\nu)$有足够快的衰减速度，窗函数就是一个局

部化的函数. 窗函数的中心 x_c 定义为：

$$x_c=\frac{(w(x),xw(x))}{(w(x),w(x))},\tag{3}$$

式中(· , ·) 表示函数 f 和 g 的内积. x_c 与 x_0 不一定相等. 窗函数的宽度则定义为

$$\Delta w=2\left[\frac{(w(x),(x-x_c)^2w(x))}{(w(x),w(x))}\right]^{1/2}.\tag{4}$$

我们注意到它是节 1.2.4 中所定义的信号空域宽度的两倍.

由于窗函数具有局部处理的功能,因此(1)式定义的变换称为短时傅里叶变换(short-time Fourier transform,简记为 STFT,以区别于常规傅里叶变换 FT). 在表达式(1)中,频率变量 ν 和坐标变量 x_0 同时出现在变换函数中,这是短时傅里叶变换和常规傅里叶变换的一个重要区别,在常规的傅里叶变换中,坐标变量和频率变量分别单独出现在信号函数 $g(x)$ 和它的频谱 $G(\nu)$ 中. 在 STFT 中,窗口宽度则隐含于 $G_w(\nu,x_0)$ 内. 正是 x_0 和窗口宽度 Δw, 使这一变换具有局部处理的功能. 改变 x_0,窗口就在空域中移动,以获取不同区域的信息,x_0 通常称为位移因子; Δw 则限制了被处理空间的范围.

与 x_c,Δw 相对应, 存在频率窗中心坐标

$$\nu_c=\frac{(W(\nu),\nu W(\nu))}{(W(\nu),W(\nu))},\tag{5}$$

和频率窗宽度

$$\Delta W=2\left[\frac{(W(\nu),(\nu-\nu_c)^2W(\nu))}{(W(\nu),W(\nu))}\right]^{1/2},\tag{6}$$

ΔW 是节 1.2.4 中所定义的信号频域宽度的两倍. 当 Δw 和 ΔW 都有限时,我们称函数 $w(x)$ 在空域和频域同时局部化.

$\Delta w\Delta W$ 称为空间-频率窗,它限制了空域和频域中被处理区域的范围. 根据 Δw,ΔW 的定义(4)、(6)及测不准关系式(节 1.2(18)式),并注意信号宽度定义的区别,我们就有

$$\Delta w\Delta W\geqslant 1/\pi.\tag{7}$$

当高斯函数取为窗函数时，(6)式中的等式成立，这种情况下短时傅里叶变换具有最小处理区域.

短时傅里叶变换的局部性，其特征在于处理过程限制在空间-频率窗内进行，且窗的位置是可变的，然而无论 Δw 还是 ΔW 都是常数，不会随信号中心频率的变化而变化，这使短时傅里叶变换在处理一些奇异性的信号时显得无力. 恰恰是在这一点上，小波变换具备比短时傅里叶变换更强的功能.

6.2.2 Gabor 变换

早在 1946 年，Gabor 就提出了下面的变换

$$\mathscr{G}_{\sigma,b}\{g\}=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}g(x)\exp(-\mathrm{i}2\pi\nu x)\exp\left[-\frac{(x-b)^2}{2\sigma^2}\right]\mathrm{d}x, \tag{8}$$

称 Gabor 变换，其中 σ 和 b 为变换的参数. 上式又可表为

$$\mathscr{G}_{\sigma,b}\{g\}=\int_{-\infty}^{\infty}g(x)\exp(-\mathrm{i}2\pi\nu x)w_\sigma(x-b)\mathrm{d}x, \tag{9}$$

式中

$$w_\sigma(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right), \tag{10}$$

正是高斯型窗函数. 因此 Gabor 变换就是高斯窗短时傅里叶变换. 窗函数中心坐标

$$x_c=0, \tag{11}$$

窗的宽度

$$\Delta w=\sqrt{2}\,\sigma. \tag{12}$$

$w_\sigma(x)$的傅里叶变换

$$W_\sigma(\nu)=\exp(-2\pi^2\sigma^2\nu^2) \tag{13}$$

也是高斯函数. 频率窗宽度

$$\Delta W=1/\sqrt{2}\,\pi\sigma, \tag{14}$$

因此有 $\Delta w\Delta W=1/\pi$.

在图 6.2 中，横坐标表示空间坐标，纵坐标表示频率，该图将空域和频域同时表达出来，称空间-频率坐标系，空-频窗则表示为图中的一个矩形. Gabor 变换空-频窗的高度和宽度都是恒定的.

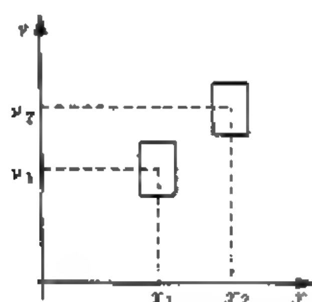


图 6.2　Gabor 变换空间-频率窗[1]

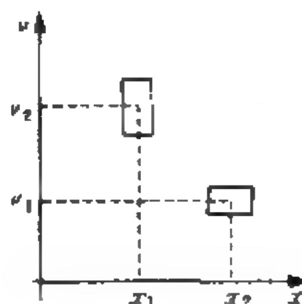


图 6.3　小波变换空间-频率窗[1]

Gabor 变换在频域中的表达式则由(2)式给出：

$$\begin{aligned}\mathscr{G}_{\sigma,b}\{g\} &= [\exp(-2\pi\sigma^2\nu^2)\exp(-\mathrm{i}2\pi\nu b)] * G(\nu) \\ &= \int_{-\infty}^{\infty} G(\nu')\exp[-2\pi^2\sigma^2(\nu'-\nu)^2]\exp[\mathrm{i}2\pi(\nu'-\nu)b]\mathrm{d}\nu' \\ &= \sqrt{2\pi}\sigma\exp(-\mathrm{i}2\pi\nu b)\mathscr{G}_{\mu,\nu}\{G\},\end{aligned} \tag{15}$$

式中

$$\mu = 1/(2\pi\sigma)^2. \tag{16}$$

可见 Gabor 变换在频域和空域中的表达式具有相似的形式.

Gabor 变换具有如下特点：

(1) 它给出一个中心位于 b，宽度为 $\sqrt{2}\sigma$ 的空间窗，从而实现空域处理的局部化；与之相应，它又给出一个中心位于 ν，宽度为 $1/(\sqrt{2}\pi\sigma)$ 的频率窗，从而实现频域处理的局部化. 用 Gabor 变换来处理信号时，处理过程限制在空-频窗内进行，窗的面积为 $1/\pi$.

(2) 由(8)式和(15)式，

$$\begin{aligned}\mathscr{G}_{\sigma,b}\{g\} &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} g(x)\mathrm{e}^{-\mathrm{i}2\pi\nu x}\exp\left[-\frac{(x-b)^2}{2\sigma^2}\right]\mathrm{d}x \\ &= \mathrm{e}^{-\mathrm{i}2\pi\nu b}\int_{-\infty}^{\infty} G(\nu')\mathrm{e}^{\mathrm{i}2\pi\nu' b}\exp[-2\pi^2\sigma^2(\nu'-\nu)^2]\mathrm{d}\nu',\end{aligned} \tag{17}$$

可看出变换是参数 σ,b 和变量 ν 的函数. 上式给出的积分是一个调制包络,载波 $\exp(-i2\pi\nu b)$ 的频率(即中心频率)ν 与参数 σ 无关,不会随 σ 的变化而变化,这正是所有短时傅里叶变换共同的缺点.

6.2.3 Morlet 小波变换

为了克服 Gabor 变换中窗口尺寸不能变动的缺点,我们可以对它略加改造. Gabor 变换的基元函数可改写为

$$
\begin{aligned}
h(x) &= \exp(-i2\pi\nu x)w_{\sigma}(x-b) \\
&= \frac{1}{\sqrt{2\pi}\sigma}\exp(-i2\pi\nu x)\exp\left[-\frac{(x-b)^2}{2\sigma^2}\right],
\end{aligned}
\tag{18}
$$

称变换的母函数. 引入参数 a,b,生成子函数

$$
\begin{aligned}
h_{a,b}(x) &= \frac{1}{\sqrt{a}}h\left(\frac{x-b}{a}\right) \\
&= \frac{1}{\sqrt{2\pi a}\sigma}\exp\left[i2\pi\left(\frac{\nu}{a}\right)(x-b)\right]\exp\left[-\frac{1}{2\sigma^2}\left(\frac{x-b}{a}\right)^2\right].
\end{aligned}
\tag{19}
$$

定义信号函数 $g(x)$ 的 Morlet 小波变换为

$$
\begin{aligned}
W_{a,b}(g) &= \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}h^*\left(\frac{x-b}{a}\right)g(x)\mathrm{d}x \\
&= \frac{e^{i2\pi\left(\frac{\nu}{a}\right)b}}{\sqrt{2\pi a}\sigma}\int_{-\infty}^{\infty}g(x)e^{-i2\pi\left(\frac{\nu}{a}\right)x}\exp\left[-\frac{1}{2\sigma^2}\left(\frac{x-b}{a}\right)^2\right]\mathrm{d}x.
\end{aligned}
\tag{20}
$$

将上式与(8)式相比较,发现 Morlet 小波变换与 Gabor 变换的实质性差别在于,小波变换的中心频率为 ν/a,随参数 a 的增大而减小. 容易算出小波变换的空间窗宽度为 $\sqrt{2}\sigma a$、频率窗宽度为 $1/\sqrt{2}\pi\sigma a$. 当中心频率增高时(a 减小),空间窗宽度变小而频率窗宽度增大,可以处理更多的高频信息;当中心频率降低时(a 增大),频率窗变小而空间窗加宽,可以容纳足够多个空间周期,以保证处理精度. 小波变换的空间-频率窗见图 6.3(见上页), 下一节将详细讨论. 图 6.4 和图 6.5 分别给出不同中心频率下 Gabor 变

换和 Morlet 小波变换的基元函数的波形. 在 Gabor 变换中,窗的宽度是常数,当中心频率增高时,一定宽度的空间窗内包含的空间周期增加,所以变换的精度是随频率而变化的;而在 Morlet 小波变换中,在处理低频信号时空间窗自动加宽,在空间窗范围内包含的信号空间周期相同,这就保证了小波变换以同样的精度去处理不同中心频率的信号,这正是小波变换与短时傅里叶变换的根本区别. 有关短时傅里叶变换的进一步研究,请参阅 Yu 和 Lu 的论文[16].

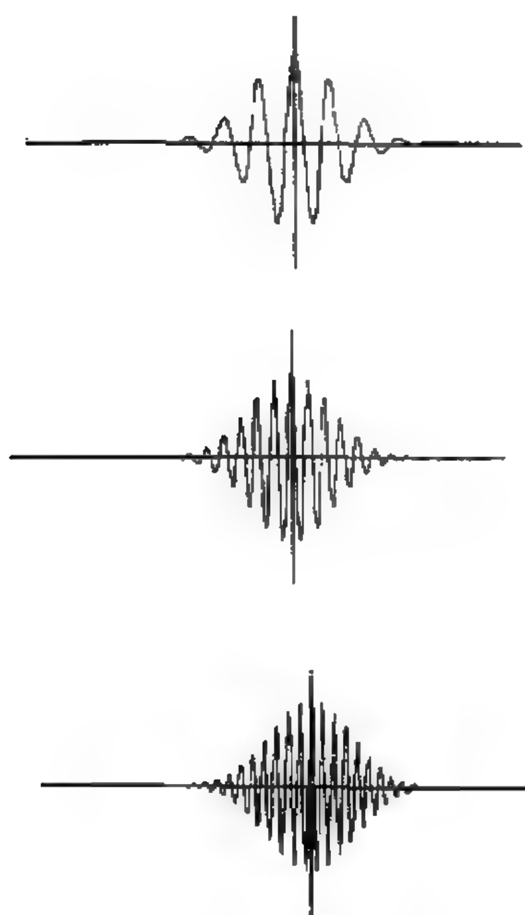


图 6.4 Gabor 变换基元函数[4]
(由上至下中心频率分别为 1,2,3)

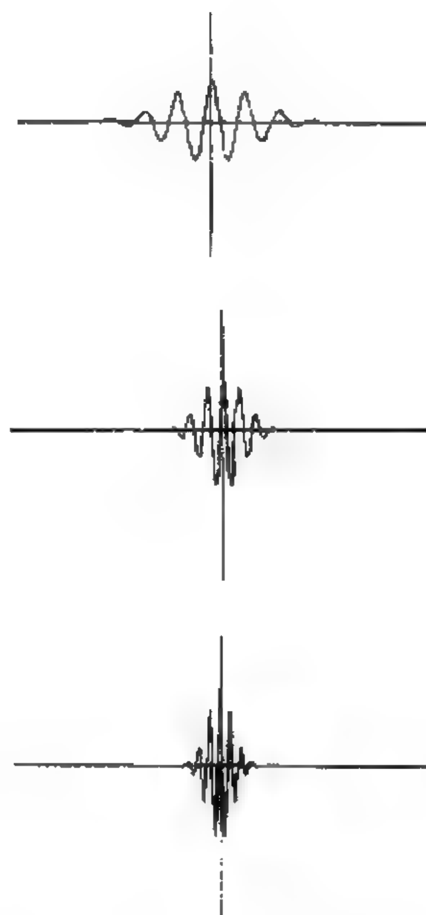


图 6.5 Morlet 小波变换基元函数[4]
(由上至下 a 分别为 1,0.5,0.33)

6.3 小波变换的定义和性质

6.3.1 小波变换的定义

母函数 $h(x)$的基本小波函数 $h_{a,b}(x)$定义为

$$h_{a,b}(x)=\frac{1}{\sqrt{a}}h\left(\frac{x-b}{a}\right), \tag{1}$$

式中 b 称为小波变换的位移因子,$a>0$ 称为伸缩因子. 由图 6.5 可以看出当 a 增大时小波的宽度加大(膨胀),而当 a 减小时小波的宽度变小(收缩). 上式表明基本小波是母函数经平移和缩放的结果. 基本小波又简称小波.

信号函数 $g(x)$的小波变换定义为小波 $h_{a,b}(x)$和 $g(x)$的内积:

$$W_{a,b}\{g(x)\}=(h_{a,b}(x),g(x))=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}h^{*}\left(\frac{x-b}{a}\right)g(x)\mathrm{d}x. \tag{2}$$

上式还可表为

$$W_{a,b}\{g(x)\}=\frac{1}{\sqrt{a}}h\left(\frac{b}{a}\right)\otimes g(b), \tag{3}$$

亦即小波变换可表为缩放后的母函数与信号函数的相关,母函数的中心位移则是相关函数的变量. 由于相关运算比较容易用光学相关器进行,因此小波变换可以用我们已熟悉的光学相关系统来实现.

并非任何函数都可以作为小波变换的函数 $h(x)$,$h(x)$必须在 $|x|\to\infty$时衰减到零. 实际使用的小波变换母函数 $h(x)$,当 $|x|\to\infty$时迅速衰减,使它的不显著为零的分量只存在于一个很小的区间内,这正是“小波”名称的来由. 实际上,也只有迅速衰减的小波才使变换(2)式具备局部化的特征.

6.3.2 小波变换在频域中的表达式

由小波变换的定义(2)式及 Parseval 定理(节 1.1(26)式)立即得到

$$W_{a,b}\{g(x)\}=\sqrt{a}\int_{-\infty}^{\infty}H^{*}(a\nu)G(\nu)\exp(\mathrm{i}2\pi\nu b)\mathrm{d}\nu, \tag{4}$$

其中 H 和 G 分别是 h 和 g 的傅里叶变换. 上式表明信号 $g(x)$的小波变换可以用 $4f$ 系统实现. 首先用第一个透镜形成输入信号 $g(x)$的傅里叶谱 $G(\nu)$,在频谱面上对 $G(\nu)$进行滤波,滤波器的表达式为小波函数 $h(x)$经缩放后的傅里叶谱函数的共轭 $H^{*}(a\nu)$,然后再经第二个透镜的傅里叶逆变换,得到 $g(x)$的小波变换. 对此我们还要详细讨论.

6.3.3 逆变换和相容性条件

小波变换(2)式的逆变换定义为

$$g(x)=\frac{1}{C_h}\int_{-\infty}^{\infty}\int_{0}^{\infty}W_{a,b}\{g(x)\}\frac{h_{a,b}(x)}{a^2}\mathrm{d}a\mathrm{d}b, \tag{5}$$

其中 C_h 满足条件

$$C_h\equiv\int_{0}^{\infty}\frac{|H(\nu)|^2}{\nu}\mathrm{d}\nu<\infty. \tag{6}$$

(6)式称"相容性条件",是逆变换存在的条件.

下面我们来证明它. $h_{a,b}(x)$可表为

$$h_{a,b}(x)=\sqrt{a}\int_{-\infty}^{\infty}H(a\nu)\exp[\mathrm{i}2\pi(x-b)]\mathrm{d}\nu, \tag{7}$$

以(4)、(7)式代入(5)式,得到

$$g(x)=\frac{1}{C_h}\int_{0}^{\infty}\frac{\mathrm{d}a}{a}\iint_{-\infty}^{\infty}H^{*}(a\nu')G(\nu')H(a\nu)$$

$$\times\exp(\mathrm{i}2\pi\nu x)\mathrm{d}\nu\mathrm{d}\nu'\int_{-\infty}^{\infty}\exp[\mathrm{i}2\pi(\nu'-\nu)b]\mathrm{d}b.$$

上式中最内部的积分为 $\delta(\nu'-\nu)$,因此

$$g(x)=\int_{-\infty}^{\infty}\left[\frac{1}{C_h}\int_0^{\infty}\frac{|H(a\nu)|^2}{a}\mathrm{d}a\right]G(\nu)\exp(\mathrm{i}2\pi\nu x)\mathrm{d}\nu. \tag{8}$$

上式成立的条件是

$$C_h=\int_0^{\infty}\frac{|H(a\nu)|^2}{a}\mathrm{d}a<\infty, \tag{9}$$

即(6)式．当 $\nu=0$ 时，相容性条件要求

$$H(0)=0, \tag{10}$$

即小波函数没有零频分量．由于

$$H(0)=\int_{-\infty}^{\infty}h(x)\mathrm{d}x=0, \tag{11}$$

意味着 $h(x)$ 必须是振荡函数，平均值为零．这样一来，可以说小波是一个满足条件(10)(或(11))的母函数 $h(x)$ 经过平移和伸缩得到的函数族 $h_{a,b}(x)$．

6.3.4 正则性[4]

从理论上讲，任何满足相容性条件的函数都可以当作小波变换的母函数，然而在实用中，为了使变换具备局部化的功能，$h(x)$ 和 $H(\nu)$ 在空域和频域中都是迅速衰减的，它们不显著为零的分量分别分布于空域和频域中的原点附近．

此外，从上一节关于 Morlet 小波变换的讨论中我们知道，变换的空间窗和频率窗的宽度随参数 a 的增大分别增大和缩小，所以我们要求 $W_{a,b}(g)$ 作为 a 的函数，应当是充分光滑的，当 $a\to0$ 时，$W\to0$，亦即要求 $W_{a,b}$ 在 $a=0$ 附近是正则的．设 $b=0$，则有

$$\begin{aligned}W_{a,b}(g(x))&=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}h^*\left(\frac{x}{a}\right)g(x)\mathrm{d}x\\&=\sqrt{a}\int_{-\infty}^{\infty}h^*(\xi)g(a\xi)\mathrm{d}\xi,\end{aligned} \tag{12}$$

将 $g(x)$ 在 $x=0$ 的邻域内展开成泰勒级数：

$$g(x)=\sum_{n=0}^{\infty}\frac{g^{(n)}(0)}{n!}x^n, \tag{13}$$

代入(12)式得到

$$W_{a,0}\{g\}=\sum_{n=0}^{\infty}\frac{g^{(n)}(0)}{n!}a^{n+1/2}\int_{-\infty}^{\infty}h^{*}(\xi)\xi^{n}d\xi$$
$$=\sum_{n=0}^{\infty}\frac{g^{(n)}(0)}{n!}a^{n+1/2}M_{n}^{*}. \tag{14}$$
式中 M_n 是小波函数 h 的 n 阶矩：
$$M_{n}=\int_{-\infty}^{\infty}h(\xi)\xi^{n}d\xi \quad (n=0,1,\cdots). \tag{15}$$

由(11)式
$$M_{0}=H(0)=0, \tag{16}$$
设
$$M_{p}=0 \quad (p=1,2,\cdots,n), \tag{17}$$
则在 0 的邻域内
$$W_{a,0}\{g\}\propto\frac{g^{(n+1)}(0)}{(n+1)!}a^{n+3/2}M_{n+1}^{*}. \tag{18}$$
随着 $a\to 0$，$W_{a,0}\to 0$ 的速率为
$$\frac{dW_{a,0}\{g\}}{da}\approx\frac{(n+3/2)}{(n+1)!}g^{(n+1)}(0)a^{n+1/2}M_{n+1}^{*}, \tag{19}$$
亦即对于一个足够平滑的函数 $g(x)$，$W_{a,0}\{g\}$以 $a^{n+1/2}$的速率随 a 趋近于零，称它为 n 阶小波函数.

由节 1.1(24)式，$M_p=0$ $(p=0,1,\cdots,n)$意味着
$$H^{(p)}(\nu)=0 \quad (p=0,1,\cdots,n),$$
表明 $\nu=0$ 是 $H(\nu)$的 n 阶零点.

此外，(16)式(即相容性条件)保证 $W_{a,0}$随 a 趋于零的速度的下限为 $\sqrt{a}$.

6.3.5 小波变换的空间-频率窗和处理过程的局部化

我们重新把小波变换的定义式写在下面：
$$W_{a,b}\{g\}=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}h^{*}\left(\frac{x-b}{a}\right)g(x)dx. \tag{20}$$
变换的基元函数 $h_{a,b}(x)$是由中心位于 x_c 的母函数 $h(x)$经位移

(位移因子为 b)、缩放(伸缩因子为 a)后形成的函数族. x_c 定义为

$$x_c=\frac{(h(x),xh(x))}{(h(x),h(x))},\tag{21}$$

空间窗宽度

$$\Delta w=2\left[\frac{(h(x),(x-x_c)^2h(x))}{(h(x),h(x))}\right]^{1/2},\tag{22}$$

这样一来,在空域中的处理就局限于下式所表示的空间窗

$$-\Delta w\leqslant\frac{x-b}{a}-x_c\leqslant\Delta w,\tag{23}$$

即

$$x\in[b+ax_c-a\Delta w,b+ax_c+a\Delta w].\tag{24}$$

在频域中,设 $H(\nu)$的中心为 ν_c,频率窗的宽度为 ΔW,ν_c 和 ΔW 分别定义为

$$\nu_c=\frac{(H(\nu),\nu H(\nu))}{(H(\nu),H(\nu))},\tag{25}$$

$$\Delta W=2\left[\frac{(H(\nu),(\nu-\nu_c)^2H(\nu))}{(H(\nu),H(\nu))}\right]^{1/2},\tag{26}$$

再引入中心位于原点的函数

$$\Psi(\nu)=H(\nu+\nu_c),\tag{27}$$

则小波变换在频域中的表达式(4)可改写为

$$W_{a,b}\{g\}=\sqrt{a}\int_{-\infty}^{\infty}\Psi^*\left[a\left(\nu-\frac{\nu_c}{a}\right)\right]e^{i2\pi\nu b}G(\nu)d\nu,\tag{28}$$

上式与(2)式对应.

因此,在频域中的处理局限在下式所表达的频率窗内进行:

$$-\Delta W\leqslant a\left(\nu-\frac{\nu_c}{a}\right)\leqslant\Delta W,\tag{29}$$

即

$$\nu\in\left[\frac{\nu_c}{a}-\frac{\Delta W}{a},\frac{\nu_c}{a}+\frac{\Delta W}{a}\right].\tag{30}$$

这样一来,空间-频率的处理就局限于下式所表示的空间-频率窗内进行:

$$[b+ax_c-a\Delta w, b+ax_c+a\Delta w]\left[\frac{\nu_c}{a}-\frac{\Delta W}{a}, \frac{\nu_c}{a}+\frac{\Delta W}{a}\right]. \quad (31)$$

这一处理过程有两个特点：

(1) 空间窗宽度 $a\Delta w$ 和频率窗宽度 $\Delta W/a$ 均随 a 的变化而变化，窗的面积 $\Delta w\Delta W$ 与 a 无关.

(2) 中心频率 ν_c/a 与带宽(即频率窗宽)之比

$$Q=\frac{\nu_c/a}{\Delta W/a}=\frac{\nu_c}{\Delta W} \quad (32)$$

与中心频率大小无关，仅取决于 $H(\nu)$. Q 是测量精度的特征量，上式表明小波变换的测量精度与频率无关.

当中心频率 ν_c/a 增大时(a 减小)，频率窗自动变宽，使小波变换作为一个检测过程，在不同频率下具有相同的精度；反之，当中心频率减小时(a 增大)，空间窗自动加宽，以容纳同样数目的信号空间周期. 有人把小波变换的这种性能比喻为“自动变焦”(zooming)，参见图 6.3.

伸缩因子 a 常称小波变换的频率变量，位移因子 b 则称为坐标变量. 下面我们讨论一个例子——实 Morlet 小波：

$$h(x)=\frac{2\cos(2\pi\nu_0 x)}{\sqrt{2\pi}\sigma}\exp(-x^2/2\sigma^2), \quad (33)$$

$$\begin{aligned} h_{a,b}(x) &= \frac{2}{\sqrt{2\pi}a\sigma}\cos\left[2\pi\nu_0\left(\frac{x-b}{a}\right)\right]\exp\left[-\frac{1}{2\sigma^2}\left(\frac{x-b}{a}\right)^2\right] \\ &= \frac{1}{\sqrt{2\pi}a\sigma}\left\{\exp\left[i2\pi\nu_0\left(\frac{x-b}{a}\right)\right]\right. \\ &\quad \left.+\exp\left[-i2\pi\nu_0\left(\frac{x-b}{a}\right)\right]\right\}\exp\left[-\frac{1}{2\sigma^2}\left(\frac{x-b}{a}\right)^2\right], \quad (34) \end{aligned}$$

它相当于上一节讨论的 Morlet 小波的实部. 容易算出

$$\left.\begin{aligned} &x_c=0, \\ &\Delta w=\sqrt{2}\,\sigma\left[\frac{1-(1-8\pi^2\sigma^2\nu_0^2)\exp(-4\pi^2\sigma^2\nu_0^2)}{1+\exp(-4\pi^2\sigma^2\nu_0^2)}\right]^{1/2}. \end{aligned}\right\} \quad (35)$$

当 ν_0 或 σ 较大时，$\Delta w\approx\sqrt{2}\,\sigma$. 在频域中，

$$H(\nu)=\exp[-2\pi\sigma^2(\nu-\nu_0)^2]+\exp[-2\pi^2\sigma^2(\nu+\nu_0)^2]. \quad (36)$$

$H(\nu)$在 $\nu=\nu_0$ 及 $\nu=-\nu_0$ 出现两个峰，关于原点对称分布，如图6.6所示.

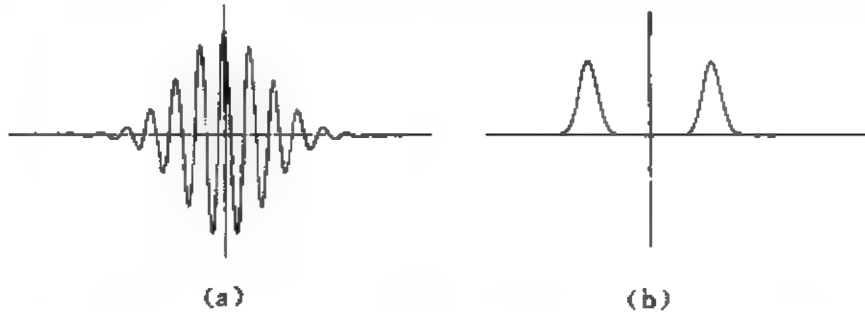


图 6.6　实 Morlet 小波 $h(x)$及它的傅里叶谱 $H(\nu)$

如果我们套用公式(25)，可算出 $\nu_c=0$. 然而从物理上看，真正起作用的是正频率，限制处理过程使频率局部化的是位于 $\nu=\nu_0$ 处的峰. 所以我们从物理的考虑出发，仅取(36)式第一项

$$H_0(\nu)=\exp[-2\pi\sigma^2(\nu-\nu_0)^2], \quad (37)$$

算出频率窗的中心位置

$$\nu_c=\nu_0 \quad (38)$$

及

$$\Delta W=1/\sqrt{2}\pi\sigma, \quad (39)$$

实 Morlet 小波变换的空间-频率窗为

$$[b-\sqrt{2}a\sigma\rho,\ b+\sqrt{2}a\sigma\rho]\left[\frac{\nu_0}{a}-\frac{1}{\sqrt{2}\pi\sigma a},\ \frac{\nu_0}{a}+\frac{1}{\sqrt{2}\pi\sigma a}\right], \quad (40)$$

式中

$$\rho=\left[\frac{1+(1-8\pi^2\sigma^2\nu_0^2)\exp(-4\pi^2\sigma^2\nu_0^2)}{1+\exp(-4\pi^2\sigma^2\nu_0^2)}\right]^{1/2}. \quad (41)$$

当 ν_0 或 σ 较大时，

$$\rho\approx1-4\pi^2\sigma^2\nu_0^2\exp(-4\pi^2\sigma^2\nu_0^2). \quad (42)$$

空间-频率窗的面积为 $\rho/\pi\approx1/\pi$，就和复数的 Morlet 小波变换结果相同.

严格说来,Morlet 小波变换不满足相容性条件的要求(10)式,因为 $H(0)\neq0$. 但当 σ 或 ν_0 较大时,$H(0)\approx0$,相容性条件是近似满足的.

6.4 实现一维小波变换的光学系统

从小波变换的定义可知,N 维信号函数的小波变换是 $2N$ 维函数,因此计算的工作量很大,尽管专门用于小波变换的超大规模集成电路已有报道,但人们仍在考虑用光学系统来实现小波变换,因为光学信息处理器具有高度的并行处理性能.

图 6.7 是一个一维小波变换光学处理系统,Li[2]和 Sheng

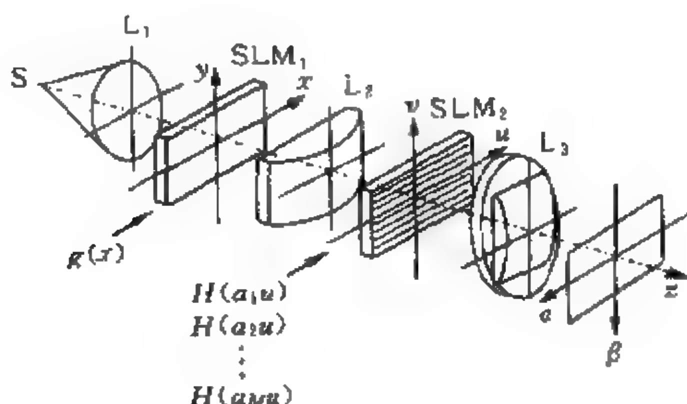


图 6.7 一维小波变换光学处理器[4]

S:相干光点光源;L_1:准直透镜;SLM_1:输入空间光调制器

L_2:柱面镜;SLM_2:谱面空间光调制器;L_3:球面、柱面复合透镜

等[4]对该系统进行了讨论. 相干光点光源 S 经准直透镜 L_1 形成平行光,照亮输入空间光调制器 SLM_1,SLM_1 的坐标系为 xy. xy 平面位于柱面透镜 L_2的前焦面,谱面 uv 则位于 L_2 的后焦面,柱面镜的母线沿 y 方向,因而 xy-L_2-uv 构成 x 方向的一维傅里叶变换器. 在 SLM_1 上输入信号 $g(x)$,经 L_2 的变换作用,在 uv 平面上形成它的傅里叶谱 $G(u)$. 在 uv 平面上放置第二个空间光调制器

SLM_2，它被分成 M 个沿 u 方向的带状区域，这些带状区域中分别显示具有不同伸缩因子 a_m 的基元函数 h 的傅里叶谱：

$$H^*(a_m u) \quad (m=1,2,\cdots,M), \tag{1}$$

假定 H 是实的，从而有

$$H^*(a_m u)=H(a_m u), \tag{2}$$

$\{H(a_m u)|m=1,2,\cdots,M\}$ 构成多通道小波变换匹配滤波器，$G(u)$ 经滤波后成为

$$H^*(a_m u)\,G(u) \quad (m=1,2,\cdots,M). \tag{3}$$

uv 平面位于球面-柱面复合透镜 L_3 前，柱面镜母线沿 u 方向，在 L_3 后面 $\alpha\beta$ 平面上放置 CCD 器件，uv-L_3-$\alpha\beta$ 构成像散系统. 在子午面(vz 平面)内，L_3 使 uv 平面成像在 $\alpha\beta$ 上，在弧矢面(uz 平面)内，柱面镜没有作用，uv 位于球面镜的前焦面，$\alpha\beta$ 位于其后焦面，构成沿 u 方向的一维傅里叶逆变换，参见图 6.8.

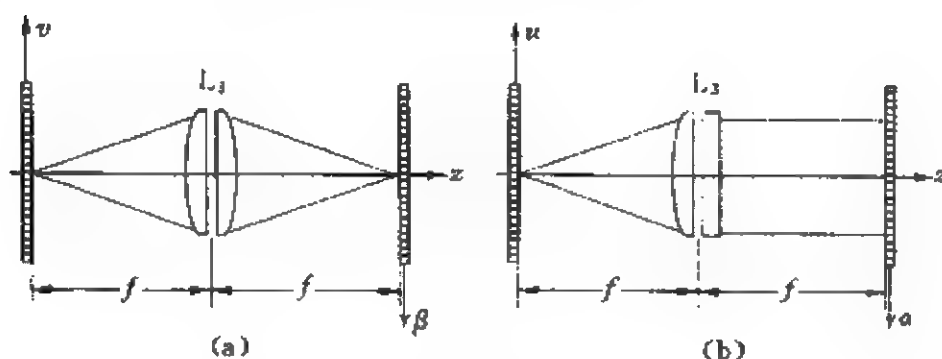


图 6.8 (a) 在子午面内构成成像系统

(b) 在弧矢面内构成一维傅里叶逆变换系统

由于成像作用，在 $\alpha\beta$ 平面上沿 α 方向相应形成 uv 平面上各带状通道的像. 对于第 m 个通道，由于沿 u 方向的傅里叶逆变换作用，得到

$$\psi(\alpha,a_m)=\int H^*(a_m u)G(u)\exp(\mathrm{i}2\pi\alpha u)\mathrm{d}u, \tag{4}$$

由节 6.3(4)式，有

$$W_{a_m,a}\{g\}=\sqrt{a_m}\psi(a,a_m)\qquad(m=1,2,\cdots,M),\qquad(5)$$

在图像处理系统中，将CCD输出的信号除以$\sqrt{a_m}$即得到小波变换$W_{a_m,a}\{g\}$.

在$W_{a_m,a}\{g\}$中，伸缩因子a是分立的，由一组M个滤波器$H(a_m u)$引入处理器；而位移因子b则是连续的，与输入平面的坐标a成正比.

6.5 用多通道匹配滤波实现二维小波变换

6.5.1 单通道小波变换系统

从节6.3(3)式我们知道，一维小波变换实际上是缩放后的变换母函数与信号函数的相关，二维小波变换也可类似定义：

$$\begin{aligned}W_{a_x,a_y,b_x,b_y}\{g(x,y)\}&=\frac{1}{\sqrt{a_xa_y}}h\left(\frac{b_x}{a_x},\frac{b_y}{a_y}\right)\otimes\psi(b_x,b_y)\\&=\frac{1}{\sqrt{a_xa_y}}\iint_{-\infty}^{\infty}h^*\left(\frac{\xi-b_x}{a_x},\frac{\eta-b_y}{a_y}\right)\psi(\xi,\eta)\mathrm{d}\xi\mathrm{d}\eta.\qquad(1)\end{aligned}$$

为了简单起见，我们令$a_x=a_y=a$，即x,y方向按相同的尺度加以缩放，得到

$$\begin{aligned}W_{a,b_x,b_y}\{g(x,y)\}&=\frac{1}{a}h\left(\frac{b_x}{a},\frac{b_y}{a}\right)\otimes\psi(b_x,b_y)\\&=\frac{1}{a}\iint_{-\infty}^{\infty}h^*\left(\frac{\xi-b_x}{a},\frac{\eta-b_y}{a}\right)\psi(\xi,\eta)\mathrm{d}\xi\mathrm{d}\eta.\qquad(2)\end{aligned}$$

在频域中，(1)式变成

$$\begin{aligned}W_{a_x,a_y,b_x,b_y}\{g\}=\sqrt{a_xa_y}\iint_{-\infty}^{\infty}&H^*(a_xu,a_yv)\Psi(u,v)\\&\times\exp[\mathrm{i}2\pi(b_xu+b_yv)]\mathrm{d}u\mathrm{d}v.\qquad(3)\end{aligned}$$

上式正是匹配滤波的频域表达式，匹配滤波可以用标准的$4f$系统实现，如图6.9所示，SLM_1，SLM_2为空间光调制器. 将二维信号函数$\psi(x,y)$经过空间光调制器SLM_1输入系统，则在L_1的频

谱面上将出现它的谱 $\Psi(u,v)$.

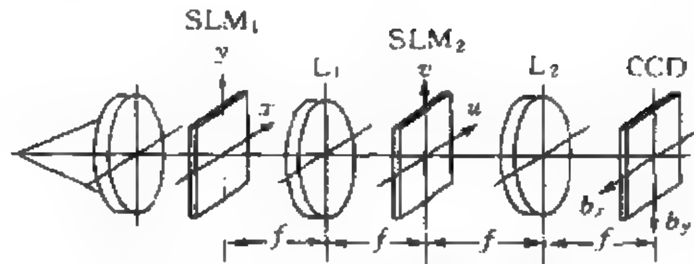


图 6.9 用 4f 系统实现小波变换

在谱面上放置第二个空间光调制器 SLM_2,将匹配滤波函数 $H^*(a_x u, a_y v)$通过 SLM_2 对 $\Psi(u, v)$进行滤波,则形成 $H^*(a_x u, a_y v)\Psi(u, v)$,再经过第二个透镜 L_2,在输出平面上得到它的傅里叶逆变换. 由(3)式可知, $H^*\Psi$ 的傅里叶逆变换,即信号 $\psi(x,y)$的小波变换 $W_{a_x,a_y,b_x,b_y}\{\psi(x,y)\}$.

我们注意到,位移因子(b_x, b_y)是与输出平面的坐标对应的变量,但伸缩因子(a_x, a_y)却是给定的,亦即我们只能对给定的伸缩因子(a_x, a_y)实现小波变换,不同的(a_x, a_y)的变换只能通过依次输入不同的匹配滤波函数 $H^*(a_x u, a_y v)$来实现,速度很慢,发挥不了光学系统并行处理的优越性.

6.5.2 用 Dammann 光栅进行多通道相关处理[9]

从节 1.1 可以知道,梳状函数 comb(x)与任意函数 $\psi(x)$相乘,等于对 $\psi(x)$的抽样; comb(x)与 $\psi(x)$的乘积再进行傅里叶变换,在谱面上得到 comb(u)与 $\Psi(u)$的卷积,相当于谱项 $\Psi(u)$在 comb(u)中各个 δ 函数的位置上重复出现,只要 $\psi(x)$的带宽是有限的,则我们总可以通过足够密集的抽样手续,使各谱项在频域内互相分离.

也就是说,对于给定的二维输入信号 $\psi(x,y)$,我们可以利用梳状函数在谱面上复制出一系列谱函数 $\Psi_{mn}(u,v)$, $m,n=0,\pm1,\pm2,\cdots$. 如果每个 Ψ_{mn}代表一个处理通道,就可以实现多通道的并行处理. 在这里,关键的器件是梳状函数器件,又称 Dammann

光栅.

一维 Dammann 光栅的透过率函数可表为

$$t(x)=\frac{1}{d}\mathrm{rect}\left(\frac{x}{\varepsilon}\right)*\mathrm{comb}\left(\frac{x}{d}\right)=\sum_{m=-\infty}^{\infty}\mathrm{rect}\left(\frac{x-md}{\varepsilon}\right), \quad (4)$$

式中 ε 是栅线宽度，d 是空间周期，见图 6.10(a).

若将 Dammann 光栅放在 4f 系统的输入面上，则在谱面上得到

$$T(u)=\varepsilon\ \mathrm{sinc}(\varepsilon u)\ \mathrm{comb}(du), \quad (5)$$

它是一个被 sinc 函数调制的梳状函数，见图 6.10(b). 若 ε 足够小，则 sinc(εu)变化缓慢，在中心附近的一些 δ 函数幅度接近于 1. 当然，ε 越小，透过的光能量就越少，这从(5)式可直接看出来.

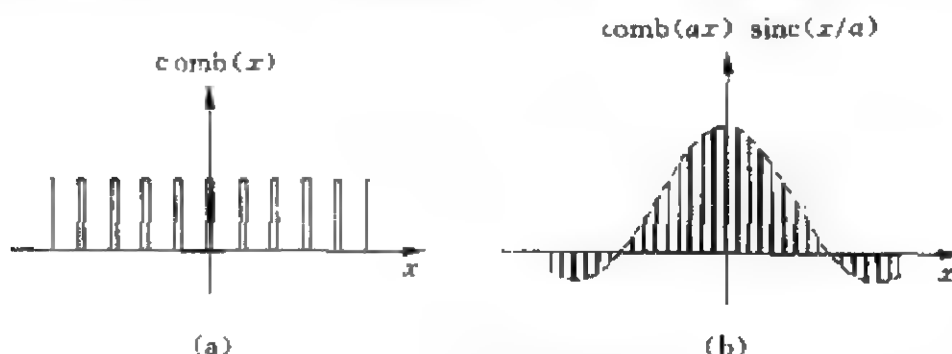


图 6.10 一维 Dammann 光栅

(a) Dammann 光栅 (b) Dammann 光栅的傅里叶谱

把两个同样规格的一维 Dammann 光栅相对旋转 90°并叠在一起，就得到二维 Dammann 光栅，其透过率函数可表为

$$t(x,y)=\frac{1}{d^2}\left[\mathrm{rect}\left(\frac{x}{\varepsilon}\right)*\mathrm{comb}\left(\frac{x}{d}\right)\right]\left[\mathrm{rect}\left(\frac{y}{\varepsilon}\right)*\mathrm{comb}\left(\frac{y}{d}\right)\right], \quad (6)$$

在 ε 很小时，上式可近似表为

$$t(x,y)=\frac{1}{d^2}\mathrm{comb}\left(\frac{x}{d},\frac{y}{d}\right), \quad (7)$$

显见二维 Dammann 光栅是一个 δ 函数点阵，它们分布在间隔为 d 的栅线的交点上，这些地方的透过率为 1，其余地方的透过率

为 0.

在图 6.9 所示的 $4f$ 系统的输入平面上放置信号 $\psi(x,y)$，并把二维 Dammann 光栅紧贴在 $\psi(x,y)$上，在光栅后的光场复振幅分布为

$$\frac{1}{d^2}\psi(x,y)\mathrm{comb}\left(\frac{x}{d},\frac{y}{d}\right). \tag{8}$$

在 uv 平面上得到它的傅里叶谱

$$\begin{aligned}\Psi(u,v)&=\Psi(u,v)*\mathrm{comb}(du,dv)\\&=\frac{1}{d^2}\sum_m\sum_n\Psi(u-mf_0,v-nf_0)=\frac{1}{d^2}\sum_m\sum_n\Psi_{mn},\end{aligned} \tag{9}$$

式中

$$f_0=1/d, \tag{10}$$

而

$$\Psi_{mn}=\Psi(u-mf_0,v-nf_0), \tag{11}$$

是一系列中心位于 $A_{mn}=(mf_0,\ nf_0)$处复现的谱项. 如设计一个匹配滤波器列阵，位于 A_{mn} 处的复数透过率为 $H_{mn}^{*}\exp[\mathrm{i}2\pi(up_m+vq_n)]$，亦即

$$F(u,v)=\sum_m\sum_n H^{*}[a_m(u-mf_0),a_n(v-nf_0)]\mathrm{e}^{-\mathrm{i}2\pi(up_m+vq_n)}. \tag{12}$$

这个滤波器中各谱项是按角度编码的，相当于用不同方向传播的平面波 $\exp[-\mathrm{i}2\pi(up_m+vq_n)]$作为参考光制成的全息匹配滤波器，通常是计算机生成的全息匹配滤波器(CGH).

经滤波后，再通过 $4f$ 系统中第二个傅里叶透镜 L_2 的作用，在输出平面上得到傅里叶逆变换

$$\begin{aligned}\psi'(b_x,b_y)&=\frac{1}{d^2}\sum_m\sum_n\iint_{-\infty}^{\infty}H^{*}[a_m(u-mf_0),a_n(v-nf_0)]\\&\quad\times\Psi(u-mf_0,v-nf_0)\mathrm{e}^{\mathrm{i}2\pi[u(b_x-p_m)+v(b_y-q_n)]}\mathrm{d}u\mathrm{d}v\\&=\frac{1}{d^2}\sum_m\sum_n C_{mn}\iint_{-\infty}^{\infty}H^{*}(a_mu,a_nv)\Psi(u,v)\end{aligned}$$

$$
\begin{aligned}
&\times e^{i2\pi[u(b_x-p_m)+v(b_y-q_n)]}\mathrm{d}u\mathrm{d}v \\
&=\frac{1}{d^2}\sum_m\sum_n C_{mn}h\left(\frac{b_x-p_m}{a_m},\frac{b_y-q_n}{a_n}\right) \\
&\quad\otimes\psi(b_x-p_m,b_y-q_n) \\
&=\frac{1}{d^2}\sum_m\sum_n C_{mn}W_{a_m,a_n,b_x-p_m,b_y-q_n}\{\psi(x,y)\},
\end{aligned}
\tag{13}
$$

式中 C_{mn} 是常数系数．亦即频域中的相移形成输出平面上的位移，以 $(p_m,\ q_n)$ 为中心形成一系列在空间相互分离的项，每一项都代表一个不同的伸缩因子 $(a_m,\ a_n)$ 的小波变换，位移因子则由以 $(p_m,\ q_n)$ 为中心的坐标来表示．这样一来，我们就用多通道 $4f$ 系统实现了二维光学小波变换，它是缩放因子的分立函数，是位移因子的连续函数．

6.5.3 用体全息存储器进行多通道相关处理[5]

重铬酸明胶（DCG）或光折变晶体（$BaTiO_3$，$LiNbO_3$，BSO 等）都可以做成体全息存储器，并可以用角度对全息图进行编码．这一性质也可以用于光学小波变换，参见图 6.11．

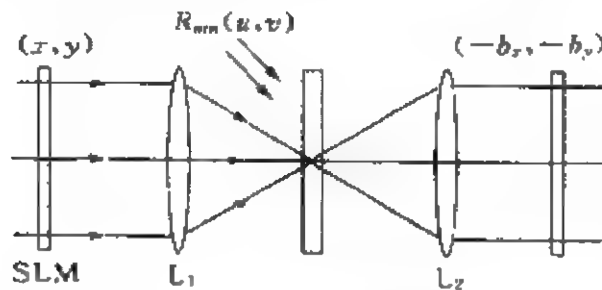


图 6.11　用光折变晶体对小波变换匹配滤波器进行编码[5]

将缩放后的母函数

$$
h_{mn}=h(x/a_m,y/a_n)
\tag{14}
$$

经 SLM_1 输入系统，则在频谱面上产生了它的傅里叶变换 $H_{mn}=a_m a_n H(a_m u,\ a_n v)$．在谱平面上放置体全息记录器件，例如光折变晶体，并用倾斜的平行参考光

$$R_{mn}=\exp[-i2\pi(up_m+vq_n)] \tag{15}$$

照射，与物光 $a_m a_n H(a_m u, a_n v)$ 相干叠加，形成全息图. 一系列 $\{h_{mn}\}$ 依次输入系统，并用不同角度的参考光 $\{R_{mn}\}$ 编码，最后得到

$$\begin{aligned} F(u,v) &= \sum_m \sum_n |a_m a_n H^*_{mn} + R_{mn}|^2 \\ &= \sum_m \sum_n \{1 + a_m^2 a_n^2 |H_{mn}|^2 + a_{mn} H^*_{mn} e^{-i2\pi(up_m+vq_n)} \\ &\quad + a_{mn} H_{mn} e^{i2\pi(up_m+vq_n)}\}. \end{aligned} \tag{16}$$

式中第三项正是我们所需的小波变换匹配滤波函数.

当我们将所分析的信号 $\psi(x,y)$ 通过 SLM_1 输入系统，它的谱 $\Psi(u,v)$ 经过 $F(u,v)$ 的滤波后得到

$$\sum_m \sum_n a_m a_n H^*(a_m u + a_n v)\Psi(u,v) e^{-i2\pi(up_m+vq_n)}, \tag{17}$$

我们略去了与匹配滤波无关的项. 再经 L_2 的傅里叶逆变换，在输出平面上得到

$$\begin{aligned} \psi'(b_x,b_y) &= \sum_m \sum_n a_m a_n \iint_{-\infty}^{\infty} H^*(a_m u, a_n v)\Psi(u,v) \\ &\quad \times \exp\{i2\pi[u(b_x-p_m)+v(b_y-q_n)]\} du dv \\ &= \sum_m \sum_n h\left(\frac{b_x-p_m}{a_m}, \frac{b_y-q_n}{a_n}\right) \otimes \psi(b_x-p_m, b_y-q_n) \\ &= \sum_m \sum_n W_{a_m,a_n,b_x-p_m,b_y-q_n}\{\psi\}. \end{aligned} \tag{18}$$

我们同样得到以 (p_m, q_n) 为中心的一系列空间互相分离的小波变换.

(16)式前二项经逆变换重合于原点，但第四项则是以 $(-p_m, -q_n)$ 为中心的卷积项，在输出平面上与相关项关于原点对称，可称为“鬼像”(ghost image)，因此在角度编码时要避免这些鬼像对其他小波变换项的干扰.

目前全息存储密度已做得相当高，因此在晶体中可以存储相当多个按方向编码的小波变换匹配滤波器.

6.6 光学小波变换匹配滤波器在图像识别中的应用

所谓图像识别或特征识别，指的是从大量信息或背景中检测某一特定的图形或指定的特征信息，并排斥其他图形信息. 在一般情况下，我们只知道需要检测的图形的特征. 对于其他图形的特征，我们事先可能并不知道，或知之甚少，但检测系统必须排斥这些图形，图像识别系统的这种性能称为"排他性". 光学小波变换识别系统在这方面的性能比常规的相关识别系统更强，因此有可能利用这一效应设计成有应用价值的小波变换图像识别系统.

6.6.1 边缘增强效应

图形的重要特征之一是它的形状或轮廓. 为了识别某一特定的图形，往往只需认定它的轮廓，而并不需要研究它的内部细节. 轮廓就是图形的边缘，一旦图形的边缘被清晰地勾画出来，这一图形就容易识别了. 相对于图形整体而言，边缘显然是局部，因此我们可以期望小波变换在边缘检测中有特殊的功效. Roberge 等详细讨论了这一课题[7].

选择"墨西哥帽"式母函数，定义为

$$h(x,y)=\frac{1}{2\pi}[2-(x^2+y^2)]\exp\left(-\frac{x^2+y^2}{2}\right), \tag{1}$$

它是高斯函数和二次函数的积. 引入

$$g(x,y)=\exp\left(-\frac{x^2+y^2}{2}\right), \tag{2}$$

则有

$$h(x,y)=-\frac{1}{2\pi}\nabla^2 g(x,y), \tag{3}$$

由节 1.1 关于 ∇^2 的傅里叶变换公式(18)，得到

$$H(u,v)=2\pi(u^2+v^2)G(u,v), \tag{4}$$

式中 G 是 g 的傅里叶变换，

$$G(u,v)=2\pi\exp[-2\pi(u^2+v^2)],\tag{5}$$

代入(4)得到

$$H(u,v)=4\pi^2(u^2+v^2)\exp[-2\pi(u^2+v^2)].\tag{6}$$

图 6.12 给出 $a=0.6$ 及 $a=0.3$ 的墨西哥帽小波函数及它们的傅里叶谱.

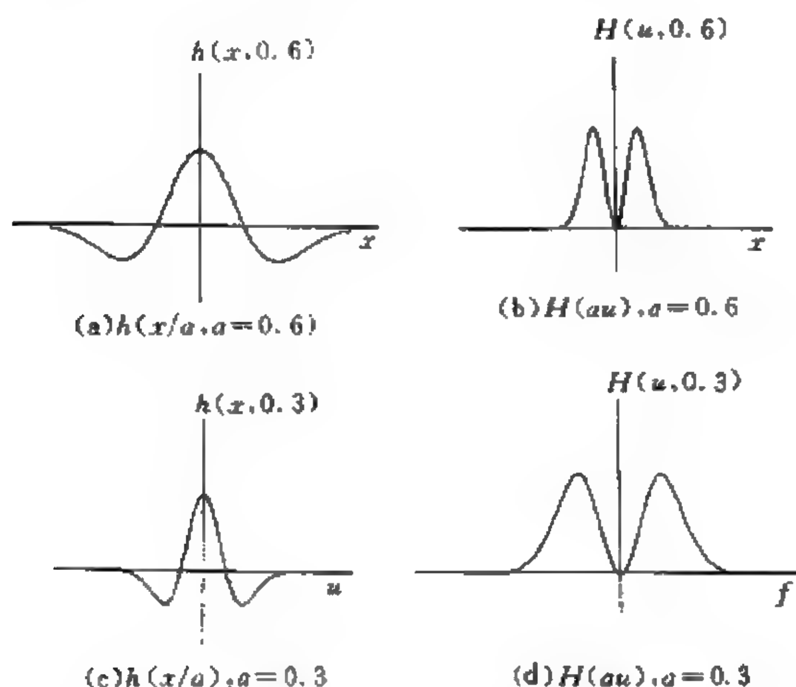


图 6.12 一维墨西哥帽小波函数及其傅里叶谱

由节 6.3 (3)式，函数 $\psi(x,y)$ 的小波变换

$$\begin{aligned}W_{a,x,y}\{\psi\}&=\frac{1}{a}h\left(\frac{x}{a},\frac{y}{a}\right)\otimes\psi(x,y)\\&=\frac{1}{a}\iint_{-\infty}^{\infty}h^*\left(\frac{\xi-x}{a},\frac{\eta-y}{a}\right)\psi(\xi,\eta)\mathrm{d}\xi\mathrm{d}\eta\\&=-\frac{1}{2\pi a}\iint_{-\infty}^{\infty}\nabla^2g\left(\frac{\xi-x}{a},\frac{\eta-y}{a}\right)\psi(\xi,\eta)\mathrm{d}\xi\mathrm{d}\eta\\&=-\frac{1}{2\pi a}\nabla^2\left[g\left(\frac{x}{a},\frac{y}{a}\right)\otimes\psi(\xi,\eta)\right],\end{aligned}\tag{7}$$

亦即 ψ 的小波变换是 ψ 和缩放后的高斯函数 g 的相关的二次导

数. 由于

$$g\left(\frac{x}{a},\frac{y}{a}\right)=\exp\left(-\frac{x^2+y^2}{2a^2}\right), \tag{8}$$

伸缩因子 a 正是高斯函数的特征尺度. g 和 ψ 相关的结果则是平滑效应,ψ 中比 a 小得多的精细结构都被平滑掉. 此外, ψ 和 g 的相关运算的结果再二次求导,$\psi(x,y)$中振幅不变的区域(常数项)及线性变化的区域(一次项)都等于零,而在振幅变化的拐点两侧不为 0. 边界正是这样的拐点.

图 6.13 为一个一维边界示意图,边界用函数 $\psi(x)$表示,图 6.13(a)同时给出 $\psi(x)$和高斯函数 $g(x/a)$; 图(b) 给出它们的相关. 相关运算对 $\psi(x)$起到了平滑的作用,结果噪声都被平均掉; 图(c) 给出小波变换作为位移因子的函数(差一个常数因子),即

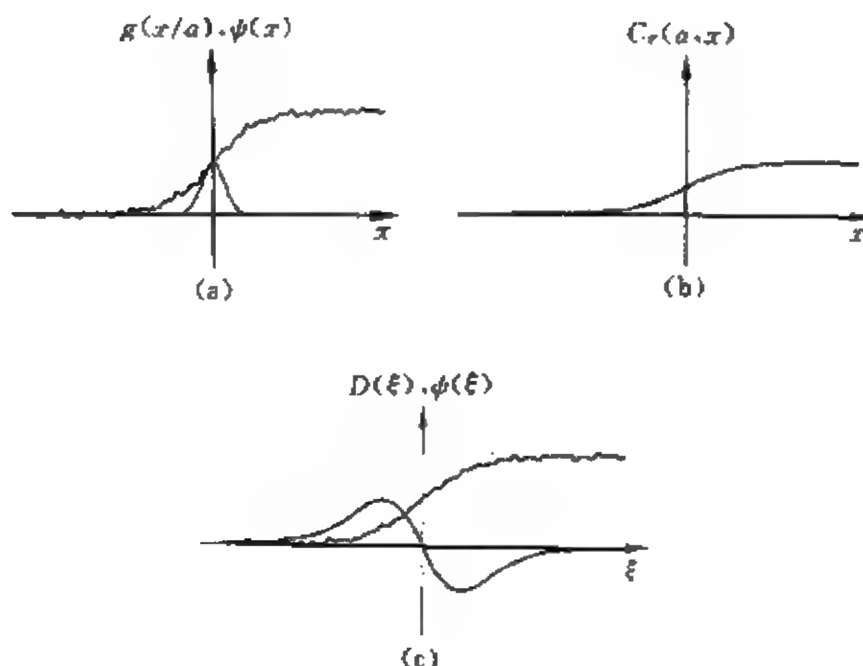


图 6.13 用小波变换来实现边界(轮廓)突出

(a) 边界函数 $\psi(x)$和高斯函数 $g(x/a)$ (b) $Cr(a,x)=g(x/a)\otimes\psi(x)$

(c)小波变换函数曲线(图中用 $D(\xi)$表示)

$$\nabla^2[g(x/a)\otimes\psi(x)], \tag{9}$$

可以在边界内外侧看到小波变换的一对正、负峰，它们明确指示了边界的位置. 边界内、外区域内小波变换函数都是零. 其最终效果恰恰是边界突出或轮廓突出，这正是我们所期望的.

在二维小波变换的情况下，伸缩因子 $\boldsymbol{a}$ 是矢量而不是标量，$\boldsymbol{a}=(a_x,a_y)$，一般情况下 $a_x\neq a_y$，从而在两个方向上的平滑效果可分别加以控制.

6.6.2 小波变换匹配滤波

图像识别课题，一般情况下仅仅勾画出图像或区域的轮廓或边界是不够的，还要求认定输入图像中是否包含要求识别的目标. 在常规光学信息处理中，我们用匹配滤波方法达到这一目的. 在小波变换信息处理中，匹配滤波方案同样适用.

设输入图像中包含要求识别的目标图像 $\psi(x,y)$ 和另一图像 $\varphi(x,y)$，首先对它进行二维小波变换，可以采用上一节中介绍的各种方法，其结果，得到了边缘增强的小波变换谱函数 $W_{a,x,y}\{\psi\}$ 和 φ 的小波变换 $W_{a,x,y}\{\varphi\}$，然后再以 $W\{\psi\}$ 和 $W\{\varphi\}$ 作为输入信号，进入第二个光学相关识别系统，例如 $4f$ 系统，或第五章中介绍的各种实现傅里叶变换的系统. 在谱面上，$W\{\psi\}$ 和 $W\{\varphi\}$ 的频谱函数分别由以下二式给出：

$$W_{a,x,y}\{\psi\}=a\iint_{-\infty}^{\infty}\Psi(u,v)H^*(au,av)\mathrm{e}^{\mathrm{i}2\pi(ux+vy)}\mathrm{d}u\mathrm{d}v, \tag{10}$$

$$W_{a,x,y}\{\varphi\}=a\iint_{-\infty}^{\infty}\Phi(u,v)H^*(au,av)\mathrm{e}^{\mathrm{i}2\pi(ux+vy)}\mathrm{d}u\mathrm{d}v. \tag{11}$$

以上两式表明，在第二个相关识别系统中，当输入函数为 $W\{\psi\}$ 和 $W\{\varphi\}$ 时，谱面上将出现 ΨH^* 和 ΦH^*.

我们用目标 $\psi(x,y)$ 所对应的小波变换匹配滤波器

$$F(u,v)=\Psi^*(u,v)H(au,av) \tag{12}$$

进行滤波，上式中

$$H(au,av)=4\pi^2a^2(u^2+v^2)\exp[-2\pi^2a^2(u^2+v^2)], \tag{13}$$

经滤波后得到

$$|\Psi(u,v)H^*(au,av)|^2 \tag{14}$$

及

$$\Psi^*(u,v)\Phi(u,v)H(au,av)H^*(au,av), \tag{15}$$

再经过傅里叶逆变换、分别得到 $W\{\psi\}$ 的自相关

$$a\iint_{-\infty}^{\infty}|\Psi^*(u,v)H(au,av)|^2\exp[i2\pi(ux+vy)]dudv$$
$$=W_{a,x,y}\{\psi\}\otimes W_{a,x,y}\{\psi\}, \tag{16}$$

及 $W\{\psi\}$ 和 $W\{\varphi\}$ 的互相关

$$a\iint_{-\infty}^{\infty}[\Psi^*(u,v)H(au,av)][\Phi(u,v)H^*(au,av)]e^{i2\pi(ux+vy)}dudv$$
$$=W_{a,x,y}\{\psi\}\otimes W_{a,x,y}\{\varphi\}, \tag{17}$$

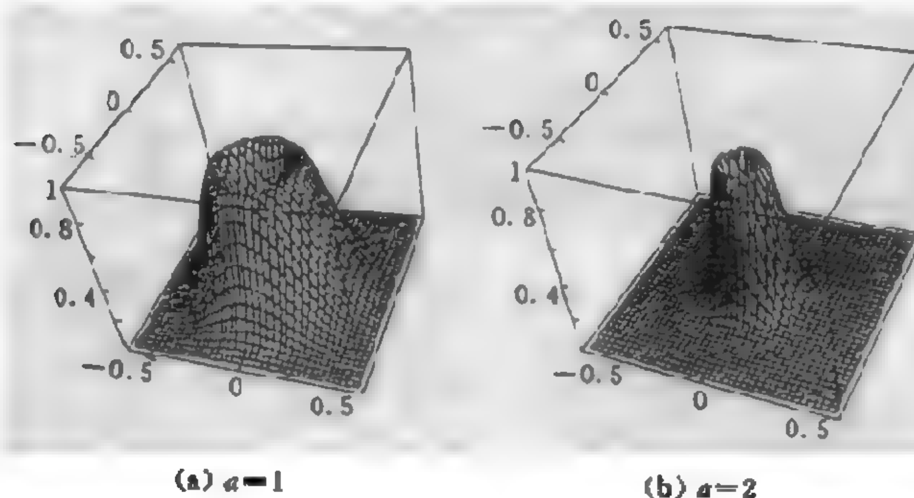
其中 $W\{\psi\}$ 的自相关给出亮斑（自相关峰），成为目标图像的特征.

由于小波变换匹配滤波方法给出经过边缘增强处理的二维小波变换函数的自相关，所以有可能获得更好的识别效果.

图 6.14(a)(b)分别给出 $a=1$ 及 $a=2$ 的墨西哥帽小波在频域中的图像. 由于小波变换匹配滤波器是 Ψ^* 与 H 的积，所以 H 正是频域中的窗函数. 当 a 增大时通带宽度减小，通带的中心频率向低频移动，可以有效地抑制高频噪声，但鉴别能力下降；反之，当 a 减小时匹配滤波器将包含更广的频段，中心频率向高频移动，将具有更高的鉴别能力，但容易受噪声的干扰，因为一般的白噪声是与带宽成正比的.

Li 等[2]给出一个用常规的匹配滤波器及小波变换匹配滤波器进行识别的结果，参见图 6.15. 图(a)是输入图像，其中右上角的字母“E”是要识别的目标；图(b)为常规的匹配滤波系统的相关输出，我们看到除了正确指示右上角目标的相关峰外，其他次峰的

高度也不低，容易造成错判；图(c)和(d)分别是用一维和二维小波变换匹配滤波器进行识别的相关输出，小波函数为墨西哥帽小波. 由于主峰明显高于次峰，信噪比远高于常规的图形识别.



(a) $a=1$　　(b) $a=2$

图 6.14　墨西哥帽式小波在频域中的图形

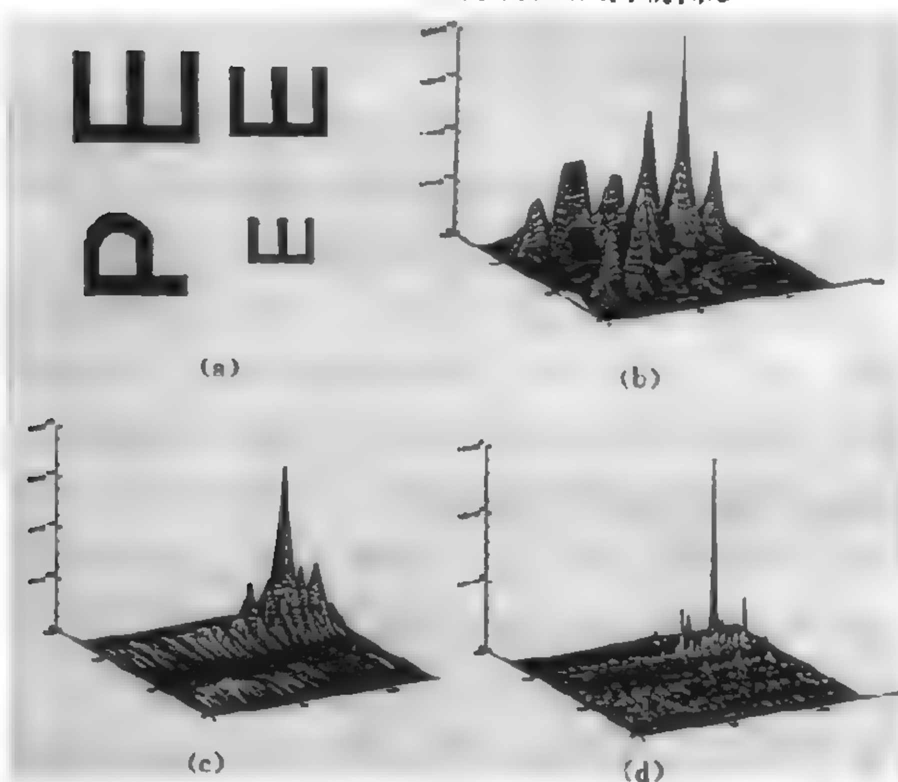


图 6.15　小波变换特征识别和常规特征识别[8]

(a) 输入图形，右上角的横写"E"为待识别目标　(b) 常规匹配滤波相关输出

(c) 一维小波变换匹配滤波相关输出　(d) 二维小波变换匹配滤波相关输出

6.7 光学 Haar 小波变换和图形边缘探测

6.7.1 Haar 变换

早在 1910 年，Haar 就首次提出了以下形式的变换函数：

$$h(x)=\mathrm{rect}\left[2\left(x-\frac{1}{4}\right)\right]-\mathrm{rect}\left[2\left(x-\frac{3}{4}\right)\right], \tag{1}$$

如图 6.16(a)为 Haar 小波函数 $h(x)$，它是以 $x=1/2$ 为中心的反对称函数，$h(x)$在区间[0,1]以外都为 0，是一个典型的“小波”，显然，该函数满足相容性条件. 此外，Haar 函数是“双极性”函数，实际上它有三个值：±1 和 0，比较容易用光学方法实现. 因此 Haar 小波变换是常用的小波变换之一. $h(x)$的傅里叶变换为

$$H(\nu)=\mathrm{i}2\mathrm{e}^{-\mathrm{i}\pi\nu}\frac{1-\cos(\pi\nu)}{\pi\nu}, \tag{2}$$

它的模如图 6.16(b)所示.

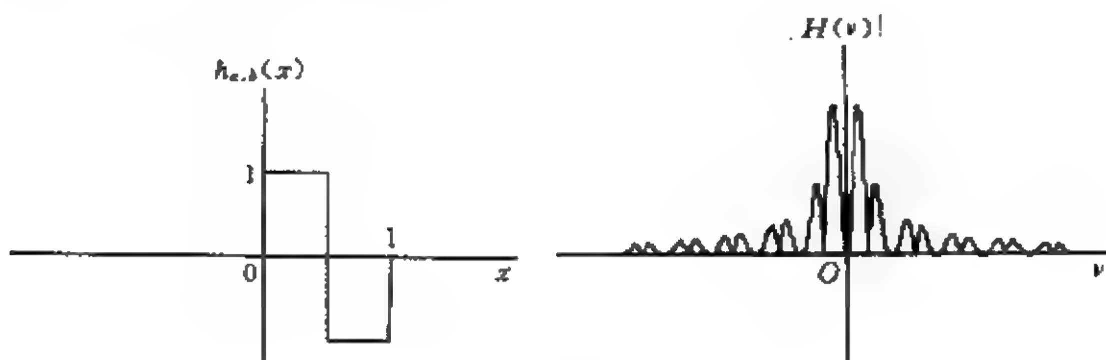


图 6.16

6.7.2 Haar 小波变换与边缘探测

Haar 小波变换是信号函数 $\psi(x)$与(1)式所定义的 Haar 函数经伸缩后的母函数 $h(x/a)$相关的结果. 对于一个给定的伸缩因子 a，Haar 小波变换的作用如下：

(1) 在小波基元函数 $h_{a,b}(x)$的正、负半周内对信号进行不加

权的积分，这事实上是一个平滑或平均的过程.

(2) 将正、负半周的积分值相减. 以上两个作用的综合结果，是在平均的意义下求差分，或求导数，恰恰是测出了图形的边缘.

图 6.17(a)(b)是对一个带有低频和高频噪声的方波进行

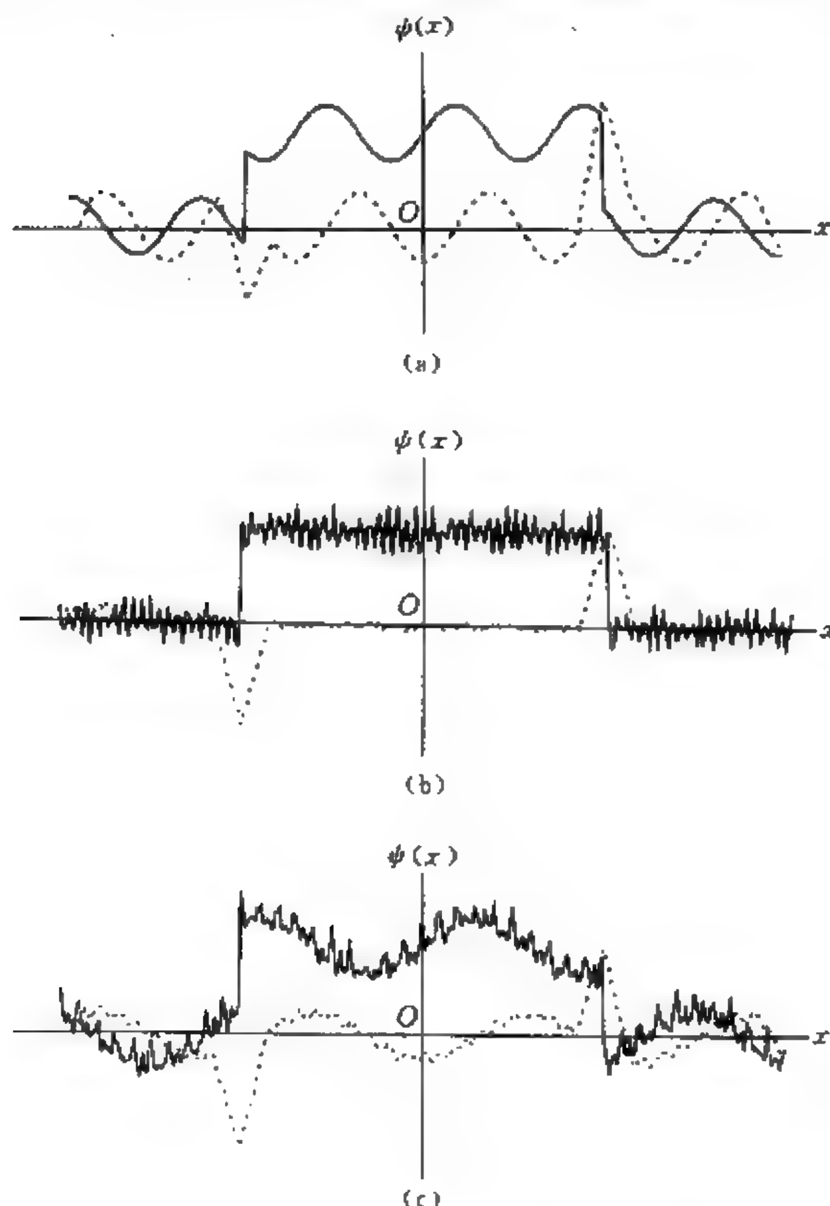


图 6.17 Haar 小波变换应用于边缘探测

Haar 小波变换的结果. 小波变换作为位移因子 b 的函数，在方波的两个边缘呈现一对峰，极值恰恰指示了边缘的位置. 图 6.17(c) 的方波同时具有低频和高频噪声干扰. 只要伸缩因子 a 选择得

当，小波变换仍然有很高的信噪比，峰很尖锐，正确的指示了边缘的所在，充分说明了 Haar 小波变换的抗干扰能力，图中实线为带有噪声的方波信号；虚线为 Haar 小波变换.

图 6.18(a)是一个带有噪声的 step 函数 $g(x)$及其 Haar 小波变换. 噪声是由空间频率为 $\nu_1=27$ 及 $\nu_2=40$ 的正弦干扰信号及随机本底噪声构成. 尽管信噪比很底，小波变换仍然正确地指示了 step 函数的边界.

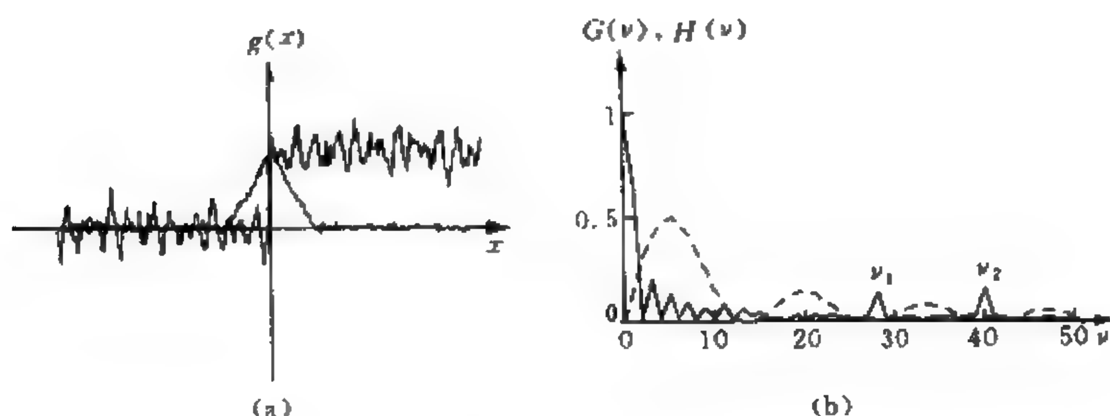


图 6.18 step 函数及其 Haar 小波变换

图 6.18(b)是信号函数和小波函数在频域中的行为，实线表示 $G(\nu)$，虚线表示 $H(\nu)$. $G(\nu)$由一个很高的零频、低频主峰及一系列旁瓣构成. 主峰主要由 step 函数的直流和低频成分贡献，而旁瓣则主要是 step 函数在边界的跃变部分贡献的.

小波函数的傅里叶谱 $H(\nu)$的主部恰恰规避了 $G(\nu)$的主峰而覆盖了主要的旁瓣. 也就是说，频率窗恰恰开在旁瓣所在的频段. 这样一来，在频域中小波变换侧重于处理旁瓣所荷载的信息；在空域中则准确指示了 step 函数的跃变，充分表现了小波变换局部处理的性能.

值得一提的是，两种干扰信号的主频 ν_1，ν_2 远在频率窗外，对变换没有贡献，而随机本底噪声的谱是宽带的，对结果的影响也不

大. 说到底，在空域中干扰信号和随机本底都是“全局”信号，小波变换的局部处理手续恰恰抑制了它们的作用，而突出了图形的局部变化——边界的贡献.

6.7.3 二维 Haar 小波变换和图形拐角测量[5]

二维 Haar 变换的母函数定义如下：

$$h(x,y)=\mathrm{rect}(x-0.5,y-0.5)+\mathrm{rect}(x+0.5,y+0.5)-\mathrm{rect}(x+0.5,y-0.5)-\mathrm{rect}(x-0.5,y+0.5), \tag{3}$$

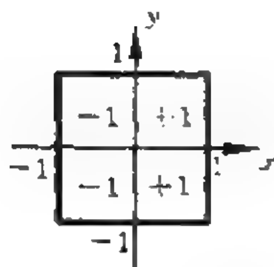


图 6.19 二维 Haar 变换母函数

如图 6.19 所示. 用 $h(x,y)$ 构作的二维 Haar 小波变换，在探测与 x 轴或与 y 轴平行的边缘时都为零，但测量任意一段既不和 x 轴平行又不和 y 轴平行的边缘时不为零，因而该变换特别适用于探测图形边缘的“拐角”.

图 6.20(a)给出一个大写字母“T”作为输入图形. 图6.20(b)是由二维 Haar 变换测量该输入图形的输出，可以看到拐角测量的效果.

二维 Haar 变换很容易改造、组合成专门进行 x 方向或 y 方向边缘测量的变换，见图 6.21. 用于测量拐角的 Haar 函数常称为“角母函数”. 用于测量边缘的 Haar 函数则称为“边母函数”. 根据图形的尺寸、边缘过渡区的宽度及噪声的频谱分布，可以采用不同的伸缩因子，以获得最佳的信息提取效果.

从一维和二维的边缘探测的讨论中，我们已看到小波变换的特点. 相对于图形整体而言，边缘(边和角)显然只是局部. Haar 小波变换对于图形内外强度不变的或缓慢变化的部分不敏感，即使对一些急剧变化的高频噪声，只要适当选取伸缩因子，Haar 小波变换仍有很强的抗干扰能力. 但在图形的边缘出现强度跃变处，小波变换却出现很强的峰，正确指示边界的位置. 可见，在边

缘探测这一类问题中，小波变换比傅里叶变换更加适用.



图 6.20 二维 Haar 小波变换应用于边、角测量[5]
(a) 输入图形
(b)(c) 小波变换输出

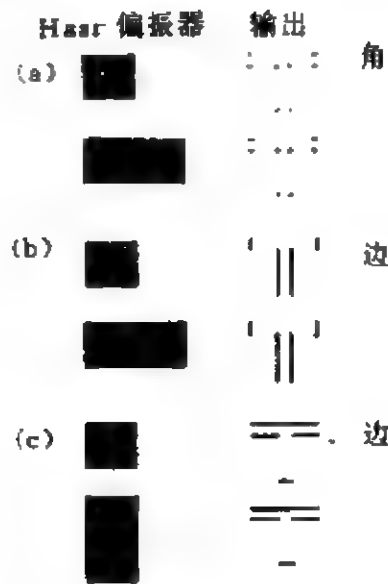


图 6.21 二维图形边、角测量的 Haar 小波变换母函数示意图[5]
(a) 角母函数 (b) y 边缘母函数
(c) x 边缘母函数

6.7.4 用投影-掩模法实现 Haar 小波变换

在频域中用匹配滤波法实现小波变换，已在上一节中详细讨论过. 本节中我们要介绍一种由 Yang 等[5]提出来的在空域中实现 Haar 小波变换的方法：投影-掩模法(shadow-casting)，参见图 6.22.

Σ_0 平面是一个点光源列阵，输出平面 Σ_3 为探测器，例如 CCD，它们分别在同样规格的透镜 L_1 和 L_2 的前、后焦面上，构成 1∶1 的成像系统；Haar 小波母函数 $h(s,t)$ 的掩模板放在 Σ_1 平面上，距 L_1 的距离为 d_1；输入信号 $g(\xi,\eta)$ 则由位于 L_1 后 d_2 处的空间光调制器 SLM 荷载. 掩模板的制作下面还要介绍. 从点光源 $P(x,y)$ 辐射的发散球面波照亮 Σ_1 上的小波函数掩模板. 在 d_1 的距离上经放大后通过 L_1 投射到 Σ_2 上. 在 L_1 和 L_2 间为平行光，没

有放大作用. 放大率即伸缩因子

$$a=f/(f-d_1). \tag{4}$$

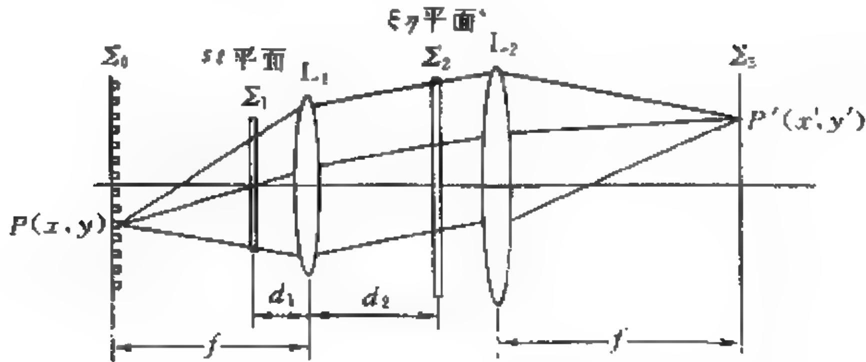


图 6.22 用投影-掩模法进行 2D-Haar 小波变换[5]

Σ_1 为小波函数掩模板；Σ_2 为输入信号 SLM；Σ_0 为点光源列阵；Σ_3 为探测器

同时,又因为 $P(x,y)$ 一般不在光轴上,使母函数发生平移. 平移量的计算可参见图 6.23. 在 xz 平面中从点 $P(x,0)$ 出发,引两条光线,一条光线通过 L_1 的中心,交 Σ_1,Σ_2 于 M,$P''(\xi,0)$,该光线不发生偏折;另一条光线经 L_1 上的 R 点发生偏折,交 Σ_2 平面于原点 O_2, 交 Σ_1 平面于 Q 点. Σ_1 平面的原点为 O_1. 光线 $PQRO_2$ 在 $\xi\eta$ 平面上的位移量为零,它在 st 平面即 Σ_1 平面上的位移量

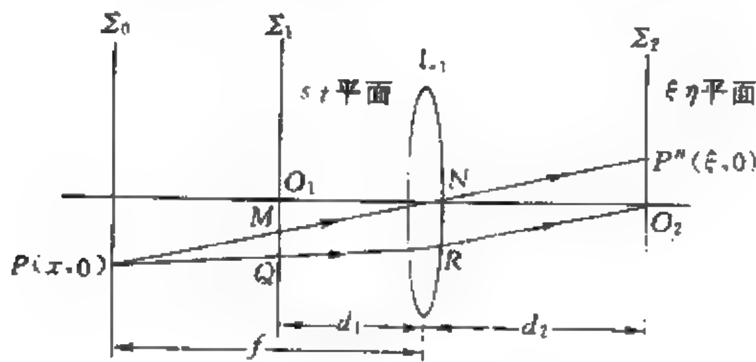


图 6.23 平移量的计算

$$b_0=O_1M+MQ=x\frac{d_1}{f}+\frac{\xi}{d}=x\left(\frac{d_1}{f}+\frac{d_2}{fa}\right), \tag{5}$$

式中 a 为伸缩因子,$MN=NR/a=\xi/a$, $NR=xd_2/f$.

b_0 放大 a 倍后换成 $\xi\eta$ 平面即 Σ_2 平面上的位移量

$$b=b_0a=x\left(\frac{d_1}{f-d_1}+\frac{d_2}{f}\right). \tag{6}$$

整个光束对于 Σ_3 上一点 $P'(x', y')$的贡献为积分

$$\psi'(b_x,b_y)=\iint h\left(\frac{\xi-b_x}{a},\frac{\eta-b_y}{a}\right)\phi(\xi,\eta)\mathrm{d}\xi\mathrm{d}\eta, \tag{7}$$

由 CCD 转换成电信号送入数据处理系统，在其中补上因子 $1/\sqrt{a}$ 即成为小波变换式

$$W_{a,b_x,b_y}\{\phi\}=\frac{1}{\sqrt{a}}\iint h\left(\frac{\xi-b_x}{a},\frac{\eta-b_y}{a}\right)\phi(\xi,\eta)\mathrm{d}\xi\mathrm{d}\eta. \tag{8}$$

倘若光源是点光源列阵，我们得到抽样的小波变换函数；若光源是均匀的连续面光源，我们得到连续分布的小波变换函数. 将掩模板沿轴向平移，可以改变伸缩因子 a 的值.

6.7.5 偏振编码[5]

掩模-投影方案中，若用非相干光源，则相干噪声不存在，信噪比很高. 但非相干光学信息处理用光强大小来表示信号，只能给出非负的分布，这显然违背了小波变换的相容性条件.

若使用相干光源，则 Haar 变换函数的双极性可以用偏振编码来实现. 设光源辐射的线偏振光沿 45°方向振动，可以用 Jones 矩阵[6,8]表为

$$\hat{J}_0=\begin{pmatrix}1\\1\end{pmatrix}, \tag{9}$$

在掩模中，+1 的部分用水平方向的偏振片构成，用 Jones 矩阵表为

$$\hat{P}_+=\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}, \tag{10}$$

−1 部分用垂直方向的偏振片构成，表为

$$\hat{P}_-=\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}. \tag{11}$$

输入的光波经过+1 和−1 部分后，分别变成

$$\hat{P}_{+}\hat{J}_0=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix}=\begin{pmatrix}1\\ 0\end{pmatrix}=\hat{J}', \tag{12}$$

及

$$\hat{P}_{-}\hat{J}_0=\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix}=\begin{pmatrix}0\\ 1\end{pmatrix}=\hat{J}'', \tag{13}$$

$\hat{J}'$和 $\hat{J}''$分别是沿 x 和沿 y 方向的线偏振光．在输入信号 SLM(即图 6.22 中的 $\xi\eta$ 平面)前放置沿−45°方向的偏振分析器，可表为

$$\hat{P}_{-\pi/4}=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}. \tag{14}$$

经过掩模中+1 及−1 部分的光波，再经过偏振分析器，变成

$$\hat{P}_{-\pi/4}\hat{J}'=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -1\end{pmatrix}=\hat{J}_{+}, \tag{15}$$

及

$$\hat{P}_{-\pi/4}\hat{J}''=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\ 1\end{pmatrix}=\hat{J}_{-}, \tag{16}$$

$\hat{J}_{+}$,$\hat{J}_{-}$分别是沿−45°及+135°方向的偏振光，它们的相位差为 π，这样一来，我们就用编码技术实现了 Haar 掩模板的制作．

参 考 文 献

[1] C. K. Chui, An Introduction to Wavelet, SanDiego, Academic, 1992.

[2] Y. Li, H. H. Szu, Y. Sheng and H. J. Caufield, Wavelet processing and optics, *Proc. IEEE*, **84**(1996), 720-732.

[3] H. H. Szu, B. Telfer and A. Lohmann, Causal analytical wavelet transform, *Opt. Eng.* **31**(1992), 1825-1829.

[4] Y. Sheng, D. Roberge and H. H. Szu, Optical wavelet transform, *Opt. Eng.* **31** (1992), 1840-1845.

[5] X. Yang, H. H. Szu, Y. Sheng and H. J. Caufield, Optical Haar wavelet transform of binary images, *Opt. Eng.* **31**(1992), 1846-1851.

[6] T. J. Burns, K. H. Fielding, S. K. Rogers, S. D. Pinski and D. W. Ruck, Optical Haar transform, *Opt. Eng.* **31**(1992), 1852-1858.

[7] D. Roberge and Y. Sheng, Optical wavelet matched filter, *Appl. Opt.* 33(1994), 5287-5293.

[8] P. G. Block, S. K. Rogers and D. W. Ruck, Optical wavelet transforms from computer-generated holography, *Appl. Opt.* **33**(1994), 5275-5278.

[9] D. Mendlovic, I. Ouzieli, I. Kiryuschev and E. Marom, Two-dimensional wavelet transform achieved by computer-generated multireference matched filter and Dammann grating, *Appl. Opt.*, **34**(1995), 8213-8219.

[10] F. T. S. Yu and G. Lu, Short-time Fourier transform and wavwlet transform with Forier-domain processing, *Appl. Opt.*, **33**(1994), 5262-5270.

[11] I. Ouzieli and D. Mendlovic, Two-dimensional wavelet processor, *Appl. Opt.*, **35**(1996), 5839-5846.

[12] 汪富泉,李后强,小波理论与分形,物理,**23**(1994),539-543.

第七章　空间光调制器

7.1　概　　论

7.1.1　空间光调制器的意义及分类

光学信息处理系统处理光波荷载的信息. 这些信息用光波的某一参数的空间分布来表征,例如强度、相位、偏振. 用光波来荷载信息有一些显著的特点如下.

首先,光波的频率高达 10^{14} Hz 以上,这就允许信号本身有很宽的带宽;其次,光波是独立传播的,两束或多束光可以在空间交叉而互不干扰,这给信息的多通道并行或交叉传播带来了极大的方便;最后,光波以并行方式传递它所荷载的信息,相对于电子系统的串行方式,光学信息处理具有大容量、高速度的特点. 最常见的透镜成像就是二维大容量信息并行处理的例子,透镜成像所涉及的是最简单的处理,就是放大或缩小,即几何光学变换.

在信息处理中,信号源(信源)和信号处理系统往往是两个独立的系统. 信源产生的信号,必须通过某种形式的接口器件,才能耦合到处理系统进行处理. 例如,常用的光学信息处理系统是相干光处理系统,它的处理能力、灵活性、多样性等都比非相干光处理系统优越. 但是大量被处理的光信号是非相干的,例如日常生活中的图形、图像,光学系统探测到的目标等. 上面我们讲过,光波是独立传播的,非相干光所荷载的信息无法直接耦合到相干光处理系统中去,通常需要一个器件,它首先接收非相干光图像,通过器件中的特殊的效应,把光学图像所表征的光强分布转换成其他物理量的二维分布(例如折射率、电压、电荷密度分布等),再通过器件的另一效应来调制处理系统中相干光的某一参数(相位、强

度、偏振等),从而完成信息从信源向处理系统的转移,以及信息由非相干光荷载向相干光荷载的变换. 实现非相干光-相干光转换的器件就是空间光调制器.

另一类常用的信号是电信号,有些信号本来就是电信号,例如各种电路产生的电信号、雷达信号、太阳辐射的射电信号等. 另有大量信号,通过传感器后转化为电信号,例如空中目标或待识别的指纹,经光学系统成像后,常常用 CCD 探测变成电信号. 日常生活中大量的图形、图像、语言、声音等,也往往通过计算机多媒体转换成视频信号. 电信号必须通过特殊设计的接口器件把电压、电流的时间变化转化为光参量的空间变化. 该接口器件也是空间光调制器.

一般来说,空间光调制器指的是这样的器件,在信源信号的控制下,它能对光波的某个参量进行调制,例如通过吸收调制振幅、通过折射率调制相位、通过偏振面的旋转调制偏振态等等,从而将信源信号所荷载的信息写进入射光波之中. 信源信号既可能是光学信号,又可能是电学信号,因此分成光寻址和电寻址的空间光调制器两大类. 当信源信号是光学信号时,我们称之为“写入光”;照射空间光调制器,并从写入光获取信息的光波称为“读出光”,因为它读出了写入信号所荷载的信息. 经空间光调制器输出的光波又称输出光,它已包含了被写入的信息. 如果按读出的方式来区分,可分成透射型和反射型两大类.

7.1.2 空间光调制器的分类及寻址方式

在光学信息处理系统中,空间光调制器是系统和外界信息交换的界面或接口. 如果按它在系统中的位置来区分,它可以用作系统的输入器件(称 input-SLM,简称 I-SLM),也可以在频谱面上作为滤波器件(称 processor-SLM, 简称 P-SLM). 原则上它也可以用于系统的输出端(称 output-SLM, 简称 O-SLM),但迄今为止,对于 O-SLM 的行为尚不清楚[1].

上面已谈过，空间光调制器是一个二维器件，也可以看成一个透过率受到写入信号控制的滤光片，表为

$$T(x,y) = T[x(t), y(t)], \tag{1}$$

$T[x(t), y(t)]$表示在时刻 t 空间光调制器在(x,y)处的复数透过率. 写入信号把信息传递到 SLM 上相应位置，以改变 SLM 的透过率分布的过程，称为“寻址”(adressing).

通常有两种寻址方式，对应于两类空间光调制器. 当写入信号是电信号(通常是视频信号或计算机的电平信号)时，采用电寻址的方法来控制 SLM 的复数透过率. 常用的电寻址的方式是通过 SLM 上两组正交的栅状电极，用逐行扫描的方法，把信号加到对应的单元上去. 电寻址又称为矩阵寻址，电寻址 SLM 记为 EA-SLM. 一对相邻的行电极和一对相邻的列电极之间的区域构成 SLM 的最小单元，又称像素(pixel)，它给出 SLM 的分辨率极限.

EA-SLM 是用得最多的空间光调制器，它将光学信息处理与近代电子技术特别是计算机-多媒体技术结合起来，构成光-电混合处理系统，应用非常广泛，本书中许多光电混合处理系统都用到 EA-SLM.

但是电寻址的 SLM 也有一些严重的缺点：首先，电信号是时间串行信号，所以电寻址是串行寻址. 而一旦在光学信息处理链中有一个 EA-SLM，二维并行串行处理就被一维串行处理代替，处理速度立即降了下来，失去了光学信息处理的重要特色.

其次，电寻址或矩阵寻址是通过条状电极来传递信息的，电极尺寸的减小有一个限度，所以像素尺寸也有限度，影响了 SLM 的分辨率. EA-SLM 的分辨率一般不高，例如磁光空间光调制器(MOSLM)的电极数(即矩阵的行×列)通常只有 256×256，液晶空间光调制器——液晶显示器件(LCD)则有 640×480 个像素单元，与电视信号的扫描频率(VGA 模式)相匹配. 由此可见 EA-SLM 对写入的空间图像进行了抽样处理. 由于抽样频率不高，图像的带宽不大，系统的空间带宽积 SW 即信息量也不够大，因为

SW 等于像素数(见节 1.2). 提高 EA-SLM 分辨率的努力一直在进行,SVGA 模式(800×600 像素)的液晶显示器件已经应用在大屏幕投影电视中,XGA 模式(1024×768)及更高分辨率的器件也在研制中,以满足高清晰度电视(HDTV)的要求.

此外,由于电极本身不透明,所以像素的有效通光面积与像素总面积之比——开口率(fill factor)较低,光能利用率不高.

近年来发展起来的数字式微反射镜器件(DMD)是一种新型的电寻址空间光调制器,它具有高效率、高对比度、多灰阶(256 个灰阶)、高色保真度等特点,具有 VGA、SVGA、XGA 及 SXGA (1280×1024)等多种规格的像素单元,与 16∶9 宽屏幕电视匹配的 2048×1152 单元的超高分辨器件也已问世. 特别是该器件是全数字化的,亦即它的灰阶、色饱和度均由数字信号控制,不仅适用于高清晰度投影电视,并符合未来的电视技术数字化趋势,称为"数字化投影技术的革命". 由于图像投影可以归属于非相干光信息处理,因此在本章中也将进行介绍.

除了矩阵寻址方式以外,EA-SLM 可以通过激光扫描等方式寻址.

当写入信号为光信号时,空间光调制器首先具备检测这一光学信号的功能,把光学信号对应的光强分布转化成电荷分布、折射率分布等等,也就是首先把光学信号写入光寻址空间光调制器中,然后由读出光通过各种效应,例如电光效应、双折射效应等,读出这一信号. 因此具有连续的寻址机构和调制机构,事实上是光探测器和光调制器的组合. 光寻址空间光调制器简写为 OA-SLM.

OA-SLM 的空间分辨率通常高于 EA-SLM. 例如液晶光阀 LCLV 的分辨率达 60 线对/mm,面积为 50 mm×50 mm,相当于 3000×3000 个像素.

OA-SLM 的最大优点在于并行寻址方式. 把写入图像成像或投影到 OA-SLM 上是在瞬间完成的,所以具有高度并行的特点. 然而高度并行并不等于高速处理,因为光探测效应的响应速度往

往不快. 具有高响应速度的光寻址空间光调制器是一个重要的研究课题.

采用光寻址时,通常 SLM 做成反射式,写入光射入 SLM 的一个端面,把信息写入 SLM,读出光射入 SLM 的另一端面,信息通过 SLM 转移到读出光中,并反射输出. 因此通常在 OA-SLM 中有一个隔离层,使读出光和写入光互不干扰. 也可以使用不同波长的光,利用滤光片消除它们之间的串扰(crosstalk).

光寻址空间光调制器常用非相干光写入,用相干光读出. 许多信号是用非相干光记录的,用非相干光写入,可以避免相干噪声,获得较高的分辨率,而用相干光读出,又可以采用相干光处理系统对信号进行处理. 上面已讲过,相干光处理技术比非相干光处理技术成熟得多.

7.1.3 常用的空间光调制器

由于空间光调制器是相干光和非相干光处理的关键器件,因此近年来国际国内开发出几十种空间光调制器,常用的有以下几种:

(一) 电寻址空间光调制器

1. 薄膜晶体管液晶显示器(thin-film-transistor liquid crystal display, TFT-LCD);

2. 磁光空间光调制器(magneto-optical SLM, MOSLM),

3. 数字微反射镜器件(digital micromirror device, DMD).

(二) 光寻址空间光调制器

1. 铁电液晶空间光调制器(ferroelectric liquid crystal SLM, FLC-SLM);

2. 液晶光阀(liquid crystal light valve, LCLV)及阴极射线管-液晶光阀(CRT-LCLV);

3. 微通道板空间光调制器(microchannel SLM,MSLM);

4. Pockels 光调制器(Pockels readout optical modalator,

PROM).

表 7.1 及 7.2 列出几种典型空间光调制器的性能.

本章介绍几种典型的空间光调制器. 有关的应用本书其他章节中都已介绍,本章只介绍一些未列入其他章节却又有重要实用价值的应用. 李育林、傅晓理[4]和赵达尊、张怀玉[5]对空间光调制器作了全面而深入的论述. 关于空间光调制器的综合评述及未来趋势的展望,请参见 Ichioka 等的文献[1],Jutamulia,Yu[24]以及 Barbier,Moddel[25]写的综述.

表 7.1 电寻址空间光调制器[1,4,19]

器件名称或代号	材料	像素数	像素尺寸/(μm×μm)	帧频/Hz	对比度	开口率
TFT-LCD(彩色)	液晶	1840×1035	40×40	60	150	0.36
TFT-LCD(彩色)	液晶	1280×1024	52×36	60	100	0.4
TFT-LCD(彩色)	液晶	1920×480	14×44	60	300	0.35
TFT-LCD(彩色)	液晶	1280×1024	34×24	60	400	0.27
Semetex SMD-256i	钇铁石榴石	256×256	56×56	30	1000	0.54
Litton MOSLM	钇铁石榴石	128×128	56×56	2000		0.54
TI-DMD(彩色)		786×576×3DMD	17×17	180	>100	0.74
TI-DMD(彩色)		2048×1152	17×17	60	>400	>0.75

表 7.2 光寻址空间光调制器[1,4]

器件名称或代号	光敏材料	有效面积 ϕ/ mm	分辨率/(lp/mm)	响应时间 t/ms	对比度
SSFLC	GaAs	25	10		
SSFLC	α-Si PC	40	28		
SSFLC	α-Si HPC		300		60
国产 LCLV,CdS	向列相液晶	40	35	30～40	100
Hughes,LCLV,Si	向列相液晶	43	12	4～16	20
Hamamatsu micro-channel	$LiNbO_3$	16	10	50～100	100
PROM	BSO	5.8	5	0.1	

注 表中 lp/mm 指每 mm 的线对数，即每 mm 的像素数，ϕ 指孔径.

7.2 磁光空间光调制器(MOSLM)

磁光空间光调制器是根据磁光效应即法拉第效应设计的，器件的基底材料为 0.5mm 厚的钆钇石榴石，它是非磁性介质. 用外延方法在它的表面生长一层掺铋的钇铁石榴石单晶薄膜，晶轴与表面垂直，它是高法拉第常数的铁磁介质膜. 然后再用光刻的方法把薄膜刻蚀成一组二维的平台阵列. 平台厚度约 5 μm，尺寸约几十 μm 见方，间距 6～12 μm. 再用集成电路技术分两次在平台的水平和垂直的间距中布上金的行、列电极栅线，在它们之间涂以绝缘层. 图 7.1 是 MOSLM 像素结构示意图. 在图中还有局部减小矫顽力的离子注入区. 图 7.2 是 MOSLM 结构的侧视图. 器件装在铝基底上，在器件周围有一个环状的预置磁饱和线圈，前后各有一片起偏器和检偏器.

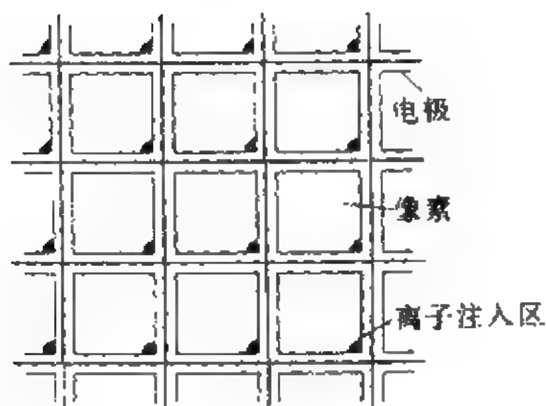


图 7.1 MOSLM 像素结构[2]

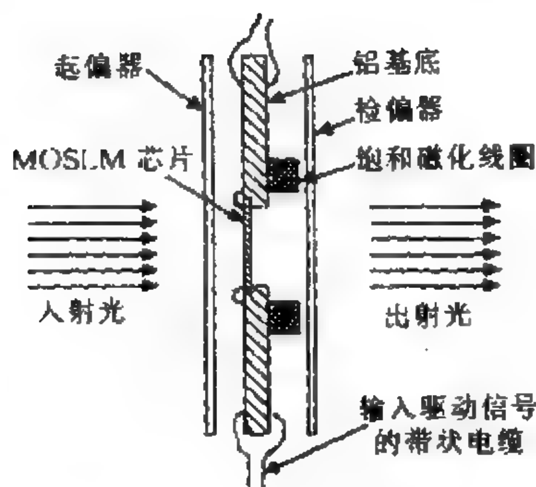


图 7.2 MOSLM 器件侧视图[4]

在使用前，先通过线圈，在 MOSLM 上加上均匀的外磁场，当撤去外磁场后，每个像素的磁性薄膜内都具有剩磁，它起到了记忆原来的外磁场方向的作用. 图 7.3 给出一对行、列电极 L_1，L_2 中的电流及它们所产生的磁场方向，A，B，C，D 为它们交点处(即寻址坐标)的四个像素单元.

设原在 MOSLM 所加的均匀外磁场，其方向从纸面向外. 加上图中所示的寻址电流后，A，C 单元中行、列电极的电流生成的磁场方向相反，正好抵消，对剩磁状态没有影响；B 单元的磁场与剩磁方向一致，也不会改变剩磁状态；只有 D 单元的外场与剩磁方向相反，若写入信号产生的磁场足够大，超过矫顽力，则 D 单元内剩磁的方向反转，即 D 单元被寻址. 而远离 L_1，L_2 交点的单元则因磁场强度太小而不起作用.

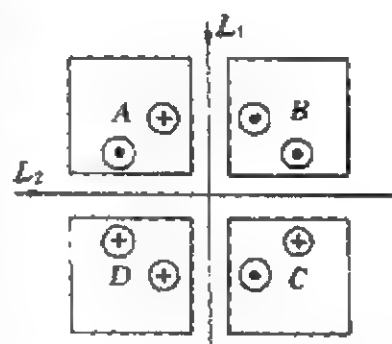


图 7.3 寻址电极中的电流及它们产生的磁场[5]

图 7.4 是 MOSLM 的工作示意图. 两个薄膜单元的剩磁磁场已被写入信号调制成相反方向，一个沿光波传播方向，另一个与之

相反. 当线偏振光沿磁光薄膜单晶的晶轴方向通过薄膜后,由于晶体中的磁光效应[6],线偏振光的振动方向分别沿顺时针和逆时针方向旋转 θ 角,它们的夹角为 2θ.

设检偏器方向与其中一个振动方向正交时,该像素即处于关态,另一像素的光强为 $I_0\sin^2(2\theta)$,这里我们未计入薄膜的吸收损失. 这样一来,MOSLM 就可以实现二元光调制,当 $\theta=45°$时反差最大.

若检偏器方向与起偏器正交,则两个像素的透过率相同,但通过它们的光波具有 180°的相位差,这种配置可以实现相位调制.

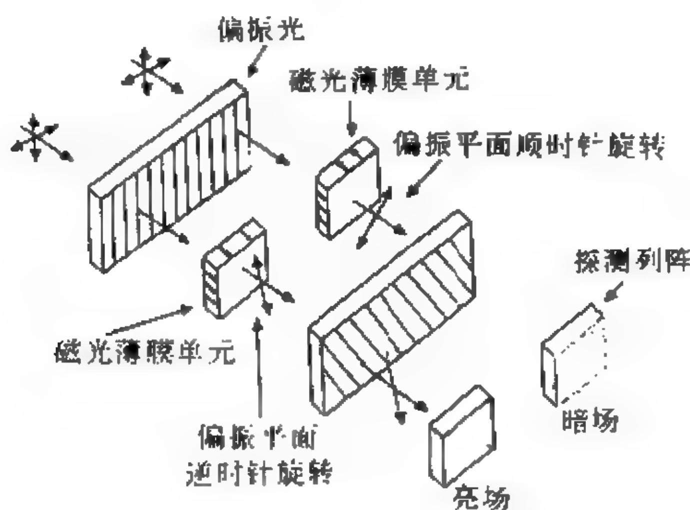


图 7.4 MOSLM 的工作示意图[2]

7.3 液晶的扭曲效应及薄膜晶体管驱动液晶显示器(TFT-LCD)

7.3.1 液晶

液晶是某些有机高分子物质在一定的条件下呈现的一种特殊的物质状态,其结构介于液体、固体之间,称为中间态,或中间相.

液晶分子一般呈长棒状,个别呈盘状、碗状,它们的分子排列介于完全规则的晶体和各向同性的液体之间. 每个液晶分子的中

心在液晶空间中的分布是随机的，但分子的取向具有有序性，亦即长棒状分子的长轴方向或盘状、碗状分子的法线方向在一定的温度范围内倾向于彼此平行，该方向称液晶分子的指向矢量方向. 三种重要的液晶结构见图 7.5.

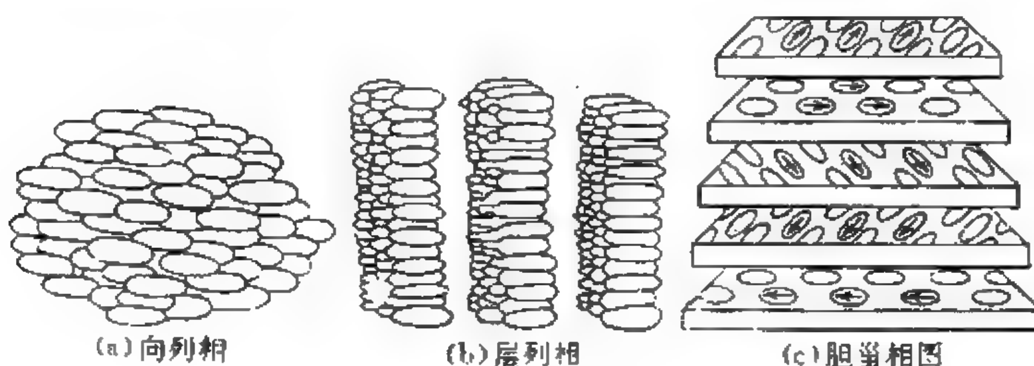


图 7.5　三种重要的液晶分子示意图[3]

因此，液晶具有双重性质，在一定程度上，既具有液体的流动性，又具有晶体所特有的各向异性. 分子结构的各向异性，必然导致电、磁、光、力学的各向异性. 此外，由于液晶分子之间的相互作用力远低于固体分子之间的相互作用力，所以液晶的各向异性在外场下会发生显著变化，这种变化远比各向异性晶体强烈. 例如 KDP 晶体的半波电压 9.3 kV，BSO 晶体的半波电压为 3.9 kV ($\lambda=632.8$ nm)，半波电压是晶体线性电光效应的一个特征参量，在振幅调制中，当外加电压达到半波电压时，调制器的透过状态从开态转成关态(参见节 7.6.1)，而表征液晶电光效应的特征参量——开关电压约为 5 V，比晶体半波电压小三个数量级，这正是液晶的流动性和各向异性双重特性的综合效果. 这一特性，使我们可以把液晶作为调制介质，构成低能耗、低电压的空间光调制器——光寻址的液晶光阀(LCLV)和电寻址的薄膜晶体管驱动液晶显示器(TFT-LCD). 尽管 TFT-LCD 是近年开发的，但作为非相干空间光调制器，它却率先投入大批量生产，并已完全商品化. 以 TFT-LCD 作为空间光调制器的计算机控制投影仪和液晶大屏幕

投影电视(LCD-projection TV，LCD-PTV)，已大批量生产，完全实用化，成为光学信息处理实用化的极个别的例子之一.

7.3.2 偏振光在扭曲介质中的传播

如果把向列相液晶放在一个经特殊处理的盒中，可以构成具有特殊的扭曲效应的液晶盒——TN 液晶盒，其结构见图 7.6.图中 ITO 为透明电极. 液晶注入一对导电玻璃之间，这两片导电玻璃的表面经过特殊处理，例如定向磨擦或倾斜蒸镀 SiO_2 导电薄膜，使它表面具有定向的结构，而且上下导电玻璃的定向结构方向正交. 向列相液晶分子为长棒状，在上面的导电玻璃 ITO 表面定向结构的作用下，接近该表面的液晶分子的指向矢趋向于 ITO 表面结构的方向排列. 而在液晶盒中，分子的排列逐渐旋转，直至下表面，正好旋转了 90°，使下表面层分子的指向矢趋向于导电玻璃 ITO 的定向结构的方向.

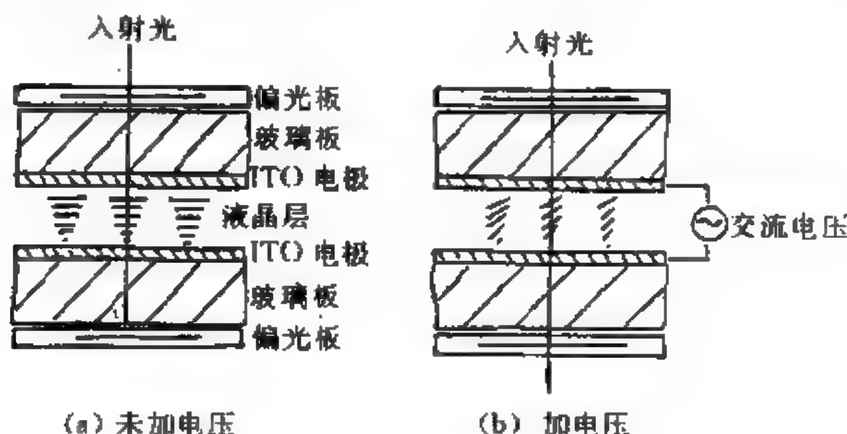


图 7.6 TN 液晶盒的结构和工作原理

如果我们把整个液晶层沿着与表面正交的 z 轴分成均匀的 N 个薄层，则每一薄层中液晶分子的长轴取向大致相同，每个这样的薄层可以看成是一个单轴晶片，分子长轴(即指向矢)的择优方向决定了光轴的方向. 在整个液晶层中，随着 z 的变化，分子长轴发生旋转，光轴也就随之旋转，形成扭曲型各向异性介质. 在线性扭曲向列型液晶中，光轴的方位 ψ 是 z 的线性函数：

$$\psi = \alpha z. \tag{1}$$

TN 盒中总的旋转角

$$\psi_d = \alpha d, \tag{2}$$

且有 $\psi_d = 90°$. TN 盒外层有一对偏振片 P,A，它们的光轴方向与上表面液晶指向矢方向相同，这一方向取为 x 轴. 设光波沿 z 方向入射，经 P 后变成沿 x 方向振动的线偏振光，它的偏振态可以用 Jones 矩阵[6]表示：

$$E_{xy} = \begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{3}$$

理论上可以证明(参见附录 A7)，由于扭曲型各向异性的作用，经液晶合透射的光波偏振态为

$$E'_{xy} = \begin{pmatrix} E'_x \\ E'_y \end{pmatrix}$$

$$= R(-\psi_d) \begin{pmatrix} \cos X - \dfrac{\mathrm{i}\Gamma \sin X}{2X} & \dfrac{\psi_d \sin X}{X} \\ -\dfrac{\psi_d \sin X}{X} & \cos X + \dfrac{\mathrm{i}\Gamma \sin X}{2X} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix}, \tag{4}$$

式中 ψ_d 由(2)式给出，为总的旋转角，$R(-\psi_d)$表示一个旋转 $-\psi_d$ 角的变换矩阵，

$$R(-\psi_d) = \begin{pmatrix} \cos\psi_d & -\sin\psi_d \\ \sin\psi_d & \cos\psi_d \end{pmatrix}, \tag{5}$$

Γ 和 X 可表为

$$\Gamma = 2\pi(n_e - n_o)d/\lambda, \tag{6}$$

$$X = \sqrt{\psi_d^2 + (\Gamma/2)^2}. \tag{7}$$

液晶是正双折射介质，$n_e - n_o \approx 0.1$，如设 $d = 25\ \mu\mathrm{m}$，$\lambda = 0.5\ \mu\mathrm{m}$，$\psi_d = \pi/2$，则 $\Gamma/\psi_d \approx 20$，从而我们有下面的近似关系：

$$X \approx \Gamma/2,$$

$$\begin{pmatrix} E'_x \\ E'_y \end{pmatrix} \approx R(-\psi_d) \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\Gamma/2} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\Gamma/2} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} = R(-\psi_d) \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\Gamma/2} \\ 0 \end{pmatrix}$$

$$= R(-\psi_d)e^{-i\Gamma/2}\begin{pmatrix}1\\0\end{pmatrix} = R(-\psi_d)e^{-i\Gamma/2}\begin{pmatrix}E_x\\E_y\end{pmatrix}, \qquad (8)$$

如将坐标轴沿 z 轴旋转 ψ_d，得到新的坐标系(ξ,η)，ξ 轴沿出射面的指向矢，则在(ξ,η)坐标系下，偏振态可表为

$$E_{\xi\eta}=\begin{pmatrix}E_\xi\\E_\eta\end{pmatrix}=R(\psi_d)E'_{xy}=R(\psi_d)R(-\psi_d)E'_{xy}=e^{-i\Gamma/2}\begin{pmatrix}1\\0\end{pmatrix}, \qquad (9)$$

可见出射光波仍然是线偏振的，振动方向沿出射面的指向矢方向，只是增加了一个不重要的相位. 条件

$$\Gamma\gg\psi_d \quad 或 \quad \alpha\ll 2\pi(n_e-n_o)/\lambda \qquad (10)$$

称弱扭曲条件，它通常都是满足的.

这样我们得到结论：在弱扭曲的情况下，当入射线偏振光的振动方向与扭曲介质表面的局部光轴一致时，振动方向将锁定在光轴的方向上，随着光轴旋转，出射光波仍是线偏振光，振动方向与扭曲介质出射表面的光轴一致. 这就是偏振光在扭曲介质中传播时的扭曲效应，这一效应常常被称为旋光效应. 尽管两种效应都表现为偏振平面的旋转，但微观机制却是根本不同的.

如果 $\psi_d=90°$，则在固定坐标系 xy 中，由(8)式得到

$$\begin{pmatrix}E'_x\\E'_y\end{pmatrix}=e^{-i\Gamma/2}\begin{pmatrix}0\\1\end{pmatrix}, \qquad (11)$$

即从液晶层射出的光波振动方向沿 y 轴. 如果检偏器光轴与 x 轴平行，则 TN 器件的透过率为零；反之若检偏器光轴沿 y 方向，则 TN 器件的透过率为最大，这两种状态分别称 TN 液晶盒的关态(暗态)和开态(亮态).

扭曲效应本来就是近似的，通常 ψ_d 并不严格等于 90°，光波在液晶中传播时，液晶中的不均匀还会引起散射. 由于上述种种原因，出射光波一般是长椭圆偏振光或部分偏振光，这些因素引起 TN 盒在暗态下的漏光，造成对比度下降.

7.3.3 扭曲向列液晶盒的工作原理

当我们在 TN 盒上通过透明电极加上纵向电压时，情况将复

杂得多.液晶的长形分子作为电偶极子,将趋向于电场方向重新排列.边界附近的分子由于边界的定向作用,受电场的影响较小,中间层的分子受电场的影响最大.这样一来,就在一定程度上抵消了扭曲效应,使器件产生一定的透过率.电压越高,趋向于电场方向排列的分子越多,扭曲效应就越弱,透过率就越高,一般 TN 盒在 5 V 的外加电压下就达到最大的透过率——开态(亮态).TN-LCD 的电光特性如图 7.7 所示.这样一来,外加的电信号就转换成 TN 盒的透过率变化,这正是电寻址液晶空间光调制器的基本原理.我们也注意到,电压和透过率的关系只在一定的范围内才近似为线性.在 TN 型液晶盒上加电压引起扭曲效应的破坏,这一效应又称混合场效应,我们将在节 7.5 中详细讨论.

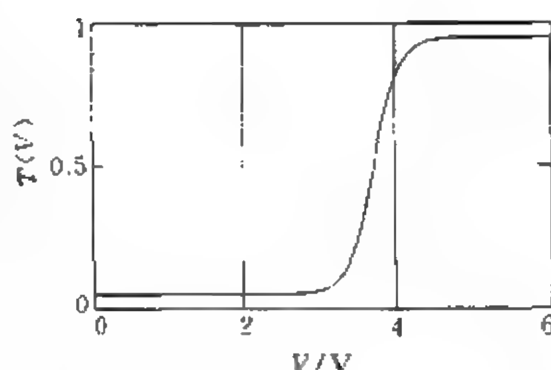


图 7.7 TN-LCD 的电光特性曲线
$T(V)$为相对透过率 V为外加电压

除此之外常用的还有超扭曲向列液晶盒(super-twinsted nematic,简写为 STN),STN-LCD 它的特性曲线比 TN-LCD 盒的更为陡峭.

7.3.4 有源矩阵驱动液晶显示器(TFT-LCD)

电寻址的空间光调制器多采用矩阵寻址的方案.通常在一块玻璃板上,形成互相绝缘的行电极和列电极,在它们的交点上用大规模集成电路技术制作薄膜晶体管 TFT.TFT 的栅极、源极和漏极分别连接行电极、列电极和显示像素.在另一块玻璃板的表面,所有像素共用一个电极,两块玻璃板之间充以扭曲型或超扭曲型液晶.图 7.8 为 TFT-LCD 的等效电路.当某一像素的行、列电极同时加上电信号时,TFT 型场效应管接通,该像素透光.顺序选

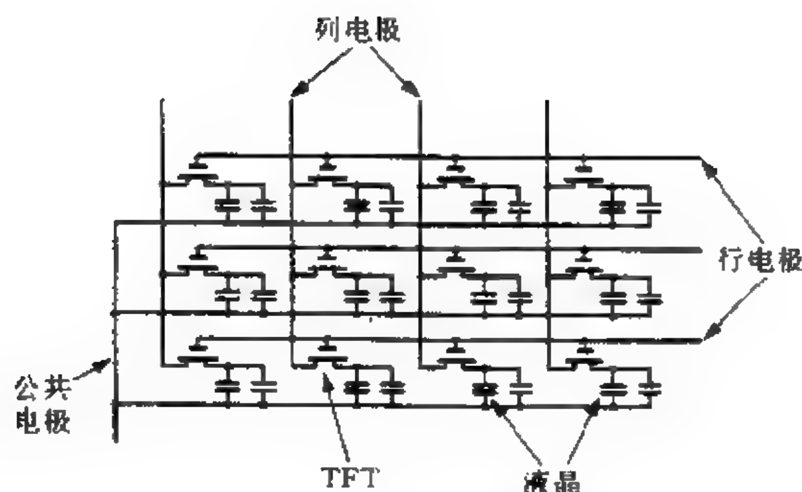
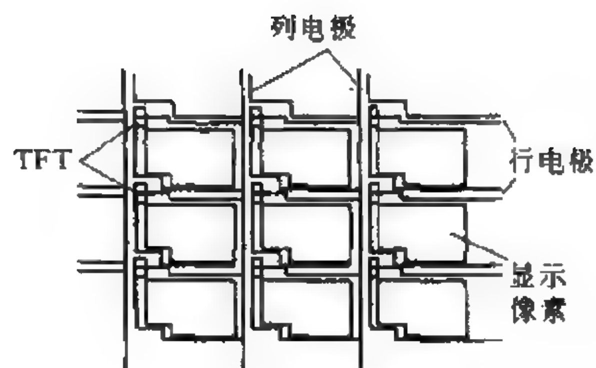


图 7.8　TFT-LCD 的等效电路

通各行电极,并同步地选通列电极,就可以控制各像素的明暗,电压的大小可控制灰阶. 图 7.9 为 TFT-LCD 的结构示意图. 在各个像素上顺序帖上红(R)、绿(G)、蓝(B)的滤光片,以一组 R,G,B 像素为一个彩色显示单元,就构成了全彩色 TFT-LCD 空间光调制器,参见图 7.10. 液晶空间光调制器不仅具有驱动电压低(<5 V)、功耗小($\mu W/cm^2$ 量级)的特点,而且由于场效应管隔离,有效地消除了相邻像素之间的干扰. 用 LCD 还可以方便地实现彩色显示. 特别是 TFT -LCD 是用大规模集成电路技术制造的,因此成品率高,成本越来越低,使它的应用越来越广泛,而广泛的应用反过来又成为 TFT-LCD 技术进步的原动力.

目前,VGA 模式(640×480 像素),3.6 英寸①(对角线尺寸)、3 英寸和 1.3 英寸的 TFT-LCD 器件已成批投产,SVGA(800×600 像素)、XVGA(1024×768 像素)的 LCD 器件也已小批投产. 美国 Xerox 公司于 1993 年演示了高达 2048×3072 像素的 LCD 样机,可显示 8.5 英寸×11 英寸幅面的图像,其分辨率已可和激光打印机媲美.[4,7]

① 1 英寸=2.54 cm.



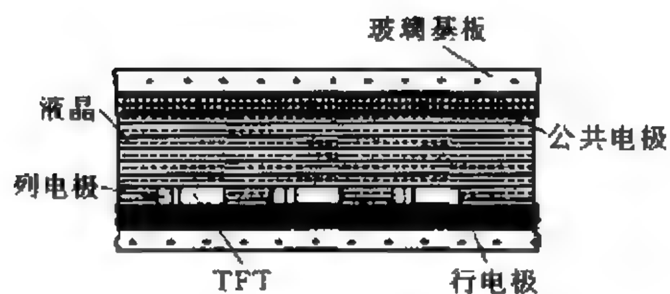


图 7.9　TFT-LCD 结构

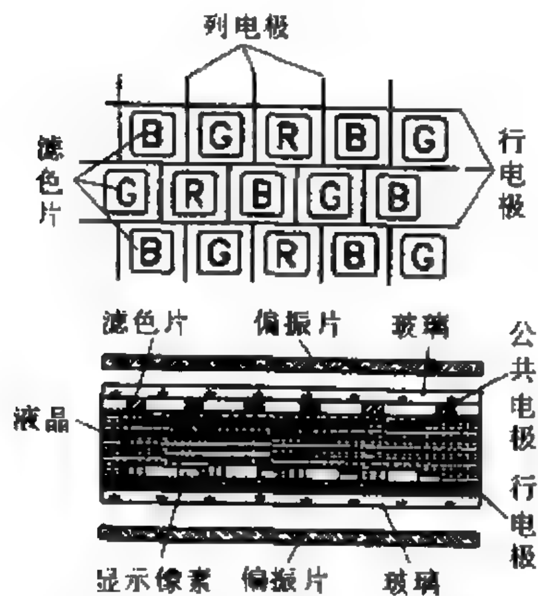


图 7.10　全彩色 TFT-LCD 液晶板

目前，TFT-LCD 的主要缺点，首先在于其透过率较低. 一对偏振片的透过率仅约 0.3，而像素的开口率(透光部分的面积与像

素总面积之比)又较低(不到0.4)，为了使LCD构成的产品(例如LCD投影电视机)的结构更紧凑，LCD的尺寸有逐步减小的趋势，目前1.3英寸对角线的LCD已成为主导产品. LCD的尺寸减小有助于降低成本，但由于电极和TFT器件的尺寸并不能成比例地减小，使小尺寸LCD的开口率更低. 耗散在器件上的光能量使器件温度升高，而温升又会引起液晶性能变差甚至永久性失效，因此使用大功率光源的LCD产品都带风冷机构. 尽管如此，LCD仍然是目前使用最为广泛的空间光调制器，而且LCD本身的性能也在一代一代地改善. 在下一节中我们将介绍使用LCD作为空间光调制器的大屏幕投影电视.

7.4 液晶显示器在非相干光信息处理中的应用——大屏幕投影电视

过去，对光学信息处理的理解局限于傅里叶光学的范畴，把那些对信号或多或少、或简或繁施行了频域处理的课题归结为光学信息处理.

广义的光学信息处理，应当包括光学信号的频域处理和空域处理，被处理的信号，既可能是相干光信号，也可能是非相干光信号或部分相干光信号. 事实上，自然界和人类生活、科学技术中大量的光信号是非相干或部分相干的. 对于非相干光信号的空域处理的历史已非常久远，例如图片的复制、放大、微缩、着色、照相负片翻正，利用“柔光镜”来改变摄影过程的全局或局部的调子(即反差度)，等等. 这些经典的非相干光信息处理有两个特点，首先，被处理的仅仅是光学图形图像，其次，处理的全过程局限于空域之中.

当各种空间光调制器开发出来以后，其他形式的信号，例如电信号、视频信号等可以通过空间光调制器输入光学系统进行处理.

一个简单而成功的例子是液晶大屏幕投影电视(LCD-projec-

tion TV，简称 LCD-PTV)，它的光学系统如图 7.11 所示．由高亮度、高色温的金属化卤钨灯 S 经非球面镜反射后形成平行光，经滤光镜 UV-IR 滤去对 LCD 有害的紫外及红外光，它们分别通过三个 TFT-LCD，在这三个空间光调制器上分别加上 R，G，B 三路视频图像信号，然后经分光镜 F_3、F_4 合成．三个聚光镜 $C_1 \sim C_3$ 及高分辨率、大视场投影物镜 L 把合成后的图像投射到屏幕上，该屏幕也可以是白色的墙壁.

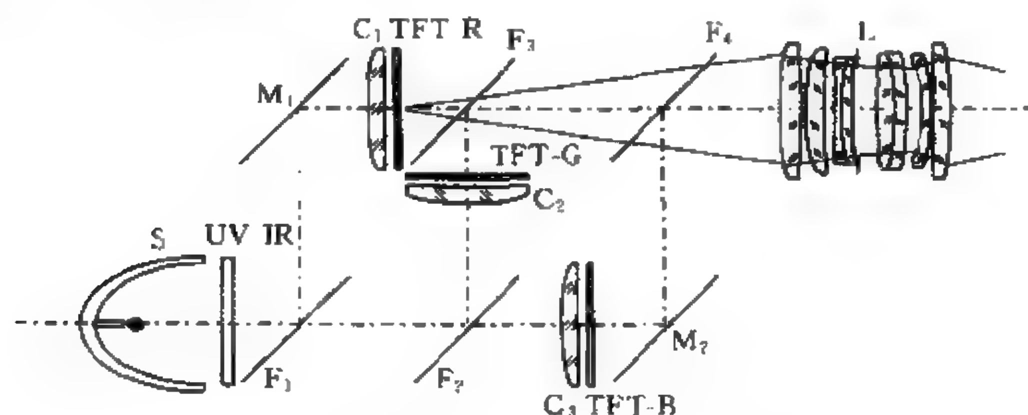


图 7.11　液晶大屏幕投影电视

S：	金属卤化钨灯	L：	投影物镜
F_1：	透蓝绿光反红光滤色片	$C_1 \sim C_3$：	聚光镜
F_2：	透蓝光反绿光滤色片	M_1，M_2：	宽波段全反镜
F_3：	透红光反绿光滤色片	UV-IR：	透可见光反红外、紫外光滤色片
F_4：	透红绿光反蓝光滤色片		

LCD-PTV 的信号既可以是电视信号，也可以是计算机、录像带或 VCD 输出的信号．它的体积、重量比常规的电视机小得多，而屏幕却可以大到 100 英寸(对角线)以上，图像清晰度、反差等均可与常规电视的相比，而且不存在大屏幕电视机的辐射危害．目前，除亮度、色饱和度比普通的电视机略微逊色外，其他性能都已赶上了普通的电视机，而屏幕尺寸则已远远超出普通的电视机.

在 LCD-PTV 中，核心器件是三个 TFT-LCD，它的作用，是用小功率电信号调制大功率光束．用于 LCD-PTV 的 TFT-LCD

的规格，有 3.6 英寸(指对角线尺寸)、3 英寸及 1.3 英寸等，1.3 英寸则是发展的趋势，使 LCD-PTV 的结构更为紧凑，重量更轻. TFT-LCD 的性能，包括像素数、反差(开态和关态透过率比)在不断的改进之中. 与此同时，高功率(例如 275 W 或更高)的金属化卤钨灯也已用在 LCD-PTV 中，使它的亮度更大，反差更高，高反差必然形成高的色饱和度，而行、列数的增加，则使 LCD-PTV 逐渐发展成为大屏幕的高清晰度电视(HDTV).

LCD-PTV 已经完全商品化，生产规模越来越大，今后，它与多媒体技术相结合，将成为家庭影院的主要设备，并广泛应用于军事指挥机关、企业、商业、科研和教育部门、娱乐场所，作为大屏幕显示、广告、计算机教学、演示和计算机辅助设计(CAD)之中.

7.5 液晶光阀

在节 7.3 中，我们介绍了矩阵寻址液晶显示器 TFT-LCD，它利用寻址电信号直接加在被寻址的液晶盒 TN-LCD 上，改变它的透过率，将扫描电信号转换成空间光的强度分布. 器件工作的基本原理，在于外加电场在一定程度上抵消了扭曲效应，从而改变了线偏振光的振动方位对扭曲介质局部光轴旋转的跟随特性. 矩阵寻址是外部电寻址方式. 如果采用光学寻址方式，则器件就是光寻址空间光调制器，通常称为液晶光阀(liguid crystal light valve，简称 LCLV)，本节将讨论这种调制器. 为了实现光寻址，器件的结构与 TFT-LCD 有所不同，但发生在液晶中的效应，本质上是完全相同的. 我们首先详细讨论存在外加电压时光波在扭曲介质中的行为，即混合场效应.

7.5.1 混合场效应

在节 7.3 中我们介绍了扭曲效应. 由于入射到 TN-LCD 盒上的线偏振光的振动平面跟随扭曲介质光轴转过 90°，出射光波

的振动平面与检偏器正交，使 TN-LCD 处于关态. 如果在 TN-LCD 盒上加纵向电压，扭曲效应将被双折射效应部分抵消，使 TN-LCD 处于开态，这一效应称为混合场效应，是美国休斯顿实验室[8,9]首先提出的. 李育林、傅晓理[4]和赵达尊、张怀玉[5]对它作了详细的阐述.

TN-LCD 的关态我们已在节 7.3.2 中详细介绍过，本节解释比较复杂的开态.

我们已讲过，液晶相处于固相和液相之间，分子间的作用力远不及固体，但比液体强. 在 TN-LCD 盒中，紧靠上下电极的液晶层由于电极表面定向结构的作用，分子的长轴方向（即指向矢方向或局部光轴方向）与表面结构的取向相同，彼此正交. 在液晶盒中，由于分子间的相互作用，各层的分子呈线性扭曲排列，在 TN-LCD 盒中共旋转 90°.

然而当我们在液晶盒上加纵向电压时，长形分子作为电偶极子，将趋向于电场排列，从而使局部光轴方向发生变化，这就是液晶的电光效应. 在电压足够高时，几乎所有的液晶分子都趋于电场方向，扭曲效应彻底破坏，液晶的光轴将转向 z 方向，从而沿 z 轴入射的线偏振光的偏振状态将不受液晶的影响，参见图 7.12.

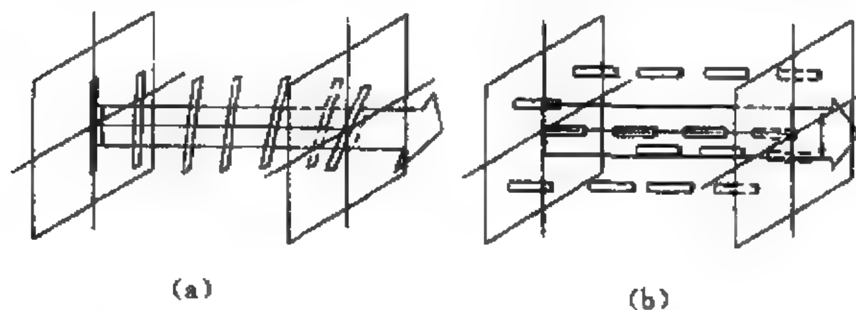


图 7.12 (a) TN-LCD 盒（向列相液晶显示器）未加电压时的扭曲效应
(b) 加电压后，液晶分子长轴沿电场方向排列，扭曲效应消失

当所加的外电压较小，尚未彻底破坏扭曲效应时，在 TN-LCD 盒中呈现出复杂的情况：液晶分子在电极定向结构的作用

力、分子间的作用力及外电场作用力的共同影响下达到平衡态. 靠近电极处，由于电极表面定向结构的影响较大，分子大体上仍沿定向结构的方向排列，扭曲效应仍占主要的地位；越接近 TN-LCD 盒中心，边界的影响越小，外电场的作用则占主要地位，长形分子趋向于外场方向(z 方向)排列，扭曲效应基本上不存在；而在边界层和中心区之间，液晶分子在 xy 平面内旋转的同时，还出现离面的分量. 仍设分子的长轴(即局部光轴)在 xy 平面的投影与 x 轴的夹角为 ψ(扭曲角)，设分子长轴与 z 轴的夹角为 β(倾斜角)，则在整个液晶层中 ψ 和 β 的分布大体上如图 7.13(a)和(b)所示，图中还用虚线表示未加电压的情况. 对于 LCLV，上下表面定向结构的夹角为 45°.

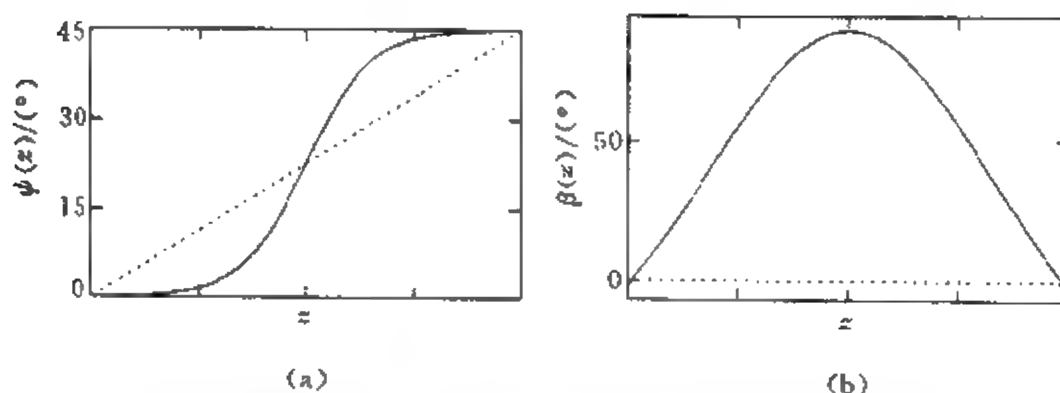


图 7.13 TN-LCD 上加较低电压时分子长轴的取向
(a) 扭曲角 ψ 作为 z 的函数 (b) 倾斜角 β 作为 z 的函数
实线为 ψ、β；虚线为 ψ'、β' 表示未加电压的情况

从图 7.13(a)可以看出，由于中心区域扭转链的松弛，使液晶分子间的作用力大大减小，因此在接近电极的边界层，液晶分子的扭曲效应比未加外电压时微弱. 在这一区段 ψ 的变化很小. 或者说在 $z=0$ 附近液晶的光轴近似沿 x 的方向，在 $z=d$ 附近近似沿 x 和 y 轴的角平分线方向，而在中心附近($z=d/2$ 处)液晶光轴倾向于 z 轴方向，ψ 的变化很剧烈，参见图 7.14. 这样一来，设在入射面附近($z\approx0$)光波的电场强度 E_0 沿 x 轴方向，与液晶光轴同向，

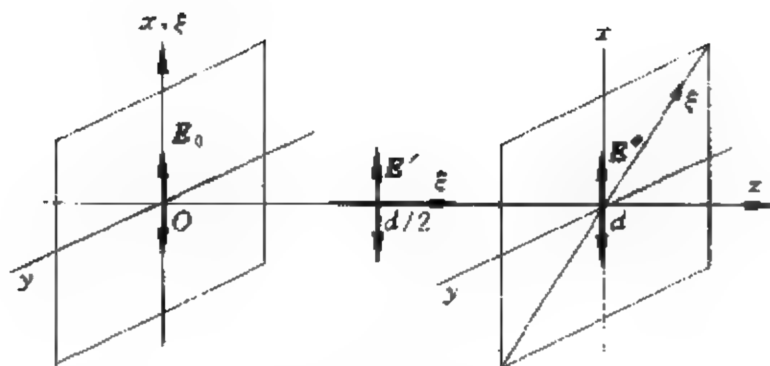


图 7.14 线偏振光在加电压后的 TN-LCD 盒中传播的近似图像；ξ 表示液晶的局部光轴，在 $z=0$ 处 ξ 与 x 轴一致，在 $z=d/2$ 处 ξ 沿 z 轴，在 $z=d$ 处 ξ 沿 x 和 y 轴的角平分线. E_0，E' 和 E'' 分别表示上述三个位置光波的电场矢量

近似为 e 光；在中心附近（$z=d/2$）由于偏振方向跟不上 ψ 和 β 的变化，近似保持不变，而光轴则趋向于 z 轴方向，因此 E' 与光轴近似垂直，为 o 光；而在 $z\approx d$ 附近 E'' 与光轴 ξ 的夹角近似为 45°，E'' 分解为 e 光和 o 光，并在这一区段获得相位差形成椭圆偏振光. 一般在 $z=d$ 处镀反光膜使光波折回，它再次通过液晶层后相位差加倍.

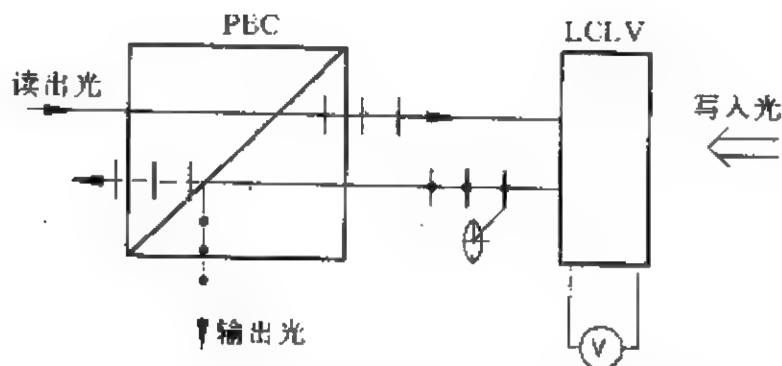


图 7.15 LCLV 的读出光路：PBC 为偏振分光棱镜. 读出光是 PBC 的 p 偏振分量，当外电压 $V=0$ 时从 LCLV 的反射光为 p 偏振分量，输出为 0；当 $V\neq 0$ 时反射光为椭圆偏振光，其中的 s 分量为输出光

图 7.15 是 LCLV 的读出光路示意图，读出光经偏振分光棱镜 PBC 后变成 p 偏振分量，振动方向与图面平行，射入 LCLV 并由器件反射．当外加电压 $V=0$ 时，反射光仍为 p 分量，完全透过偏振分光棱镜 PBC，反射的输出信号强度为 0；当 $V\neq0$ 时，反射光为椭圆偏振光，输出的 s 分量不为 0．但输出信号强度与电压不成正比，当电压 V 过大时，绝大部分液晶分子都沿 z 方向排列，$z=d$附近沿 45°方向排列的液晶层变得很薄，从而偏振光在这一区域获得的相位差反而变小．

我们同时也看到，之所以选择 $\psi_d=45°$的原因，在于当 $\psi_d=45°$ 时，双折射效应引起的相位差最大．

7.5.2 液晶光阀的结构和工作原理

图 7.16 为液晶光阀 LCLV 的结构．在读出光一侧(左侧)，液晶层外面有一层绝缘层 SiO_2，它上面有定向结构；再外层是透明电极、玻璃基底和增透介质膜层．在写入光一侧(右侧)，依次为 SiO_2 定向绝缘层、光阻层、反射介质膜层(介质镜)、CdS 光导层、透明电极和玻璃基底．其中，光阻层是 CdTe 薄膜，它的光密度(即透过率倒数的常用对数值)达 4 左右，有效地隔离了写入光和读出光．由于光阻层、介质膜反射层、绝缘层等都很薄，交流阻抗很小，外电压主要降落在液晶层和光导层上．在写入图像的暗区，光导层的电阻很大，外电压主要降落在光导层上，液晶层上的电压降很小，不足以产生明显的电光

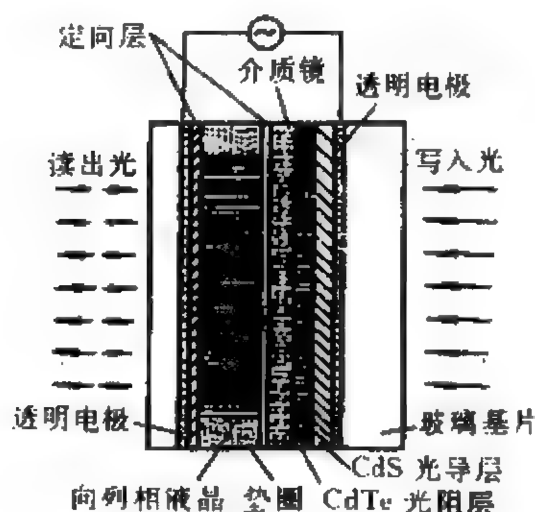


图 7.16 液晶光阀(LCLV)结构示意图[4]

效应，扭曲效应仍是主要的，光波振动平面跟随液晶介质局部光轴旋转，从器件输出的反射光仍然是 p 分量，全部透过 PBC，反射光强近似为 0；在写入图像的亮区，由于光电效应，光导层的电阻变小，它的电压降变小，液晶层上的电压降相应加大，从而引起电光效应，输出光成为椭圆偏振光，从 PBC 反射的光强不为 0. 在写入图像亮度不同的区域，输出光强也不同，写入图像通过 LCLV 的混合场效应调制了读出光.

通常写入光可以是非相干光，而读出光则为相干光，从 PBC 输出的光信号直接进入相干光处理系统.

LCLV 的反差高达 100：1，电源电压 5～10 V，分辨率高达几十线对/mm 到 100 线对/mm，通光孔径 $\phi\approx 50$ mm. 但它的响应速度不快，这是由于液晶分子在外场作用下的取向变化滞后于外场的变化，滞后时间约在 10 ms 的量级，而且器件的电容较大，作为整体的响应速度与电视图像大体匹配.

图 7.17 给出典型的 LCLV 响应曲线，横坐标为外加电压 V，纵坐标为器件输出光强与读出光强之比，即透过率 T. 可以看出，在 $V=2$～4 V之间，T 与 V 大体上呈线性关系，当 V 超过 4V 后，T 反而下降.

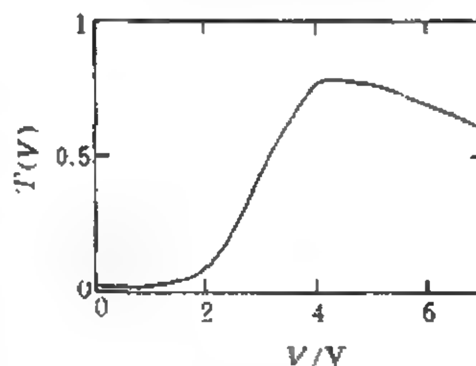


图 7.17 LCLV 透过率 T 与外加电压 V 的关系[5]

为了改进 LCLV 的性能，可以用半导体 Si 代替硫化镉 CdS，它构成的光导层暗阻抗极高，保证了关态器件不漏光，其光电灵敏度和时间响应性能也比 CdS 好[13,14]. 近年来发展起来的铁电液晶光阀(FLC)及表面稳定型铁电液晶光阀(SSFLC)，则具有更快的响应速度及记忆特性[15,16,17,18].

7.5.3 阴极射线管耦合液晶光阀(CRT-LCLV)[4]

CRT 是常用的图像输出设备，将 CRT 与 LCLV 通过光纤面板结合起来，使 CRT 的输出图像作为 LCLV 的写入信号，就构成 CRT-LCLV，其结构如图 7.18 所示. CRT 的图像通过光纤面板直接照射 LCLV 的光导层，就可以完成写入操作.

CRT-LCLV 能够实时地将计算机输出的图像转换成光学图像. 它的亮度大、分辨率高，对比度好，特别适宜于光电混合处理.

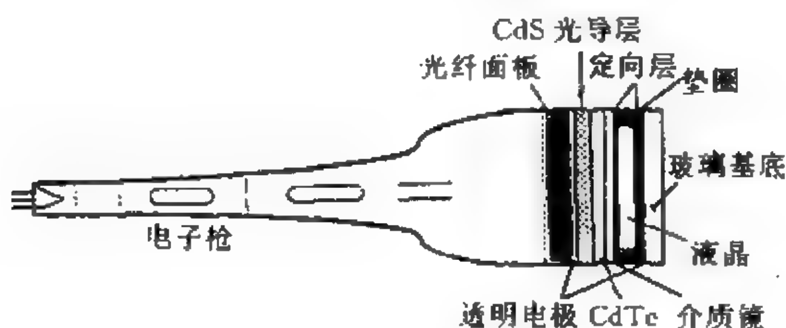


图 7.18 CRT-LCLV 结构原理图[4]

7.6 线性电光效应和 PROM 器件

7.6.1 线性电光效应[3,6]

光波是高频电磁波，但一般情况下光波的场是很弱的，当光波射入介质中时，它使介质中的带电微粒在平衡位置附近作微小的高频振动所引起的效应十分微弱，仅当用强激光照射介质时，才会引起显著的非线性效应.

然而当介质中除了光波的电磁场外同时存在其他强电场时，情况将复杂得多. 外电场将引起介质中束缚电荷分布的显著变化，从而影响介质微观结构的对称性. 例如使各向同性介质变成各向异性的，或改变晶体的对称性，导致介质光学性质的变化. 这样一来人们就有可能通过外电场和写入光波的共同作用，改变读出光波的参数分布，从而实现空间光调制. 我们首先介绍线性电

光效应即 Pockels 效应.

当介质处于恒定电场中,介质的参数将成为电场强度 $\boldsymbol{E}$ 的函数(在这一节中,$\boldsymbol{E}$,E_i 分别表示外加电场强度矢量及其分量,不要与光波的电磁场混淆起来). 在通常情况下,可将这些参数展开成 $\boldsymbol{E}$ 的分量的级数. 例如介电张量 ε 的逆张量 η 可表为

$$\eta_{ij}(\boldsymbol{E}) = \eta_{ij}(0) + \sum_{k}\gamma_{ijk}E_k + \sum_{k,l}S_{ijkl}E_kE_l + \cdots, \qquad (1)$$

式中 γ_{ijk} 称为线性电光系数或 Pockels 系数,相关的效应称线性电光效应或 Pockels 效应;S_{ijkl} 称为二次电光系数或 Kirr 系数,相应的效应称二次电光效应或 Kirr 效应. γ,S 不为零的介质称电光介质.

考虑到电场的作用后,晶体法线椭球(或折射率椭球)方程式可表为

$$\sum_{i,j}\eta_{ij}(\boldsymbol{E})\xi_i\xi_j = 1, \qquad (2)$$

式中 ξ_i,ξ_j 为直角坐标系 x,y,z 的任一分量.

可以证明[6],η_{ij} 在透明介质中是厄米型的,在非旋光介质中是实对称的,下标 i 和 j 可以对易. 为了简单起见,采用下面的缩写规则:

$$\left.\begin{aligned}&(11)\to 1,\quad (22)\to 2,\quad (33)\to 3,\\ &(23,32)\to 4,\quad (31,13)\to 5,\quad (12,21)\to 6.\end{aligned}\right\} \qquad (3)$$

这一规则使 $3\times3\times3$ 张量 γ_{ijk} 缩写成为 6×3 张量:

$$\left.\begin{aligned}&\gamma_{11k}=\gamma_{1k},\quad \gamma_{22k}=\gamma_{2k},\quad \gamma_{33k}=\gamma_{3k},\\ &\gamma_{23k}=\gamma_{32k}=\gamma_{4k},\quad \gamma_{31k}=\gamma_{13k}=\gamma_{5k},\quad \gamma_{12k}=\gamma_{21k}=\gamma_{6k},\\ &\qquad k=1,2,3\end{aligned}\right\} \qquad (4)$$

以下我们用 $\gamma_{\mu i}$ 表示缩写后的张量,$\mu=1,2,\cdots,6$;$i=1,2,3$.

外电场 $\boldsymbol{E}$ 虽然比光波的电场强,但仍然远低于原子内部的场强. 在存在线性电光效应时,二次效应即可忽略. 我们仅考虑线性电光效应,在(1)式中只保留 E_i 的线性项,得到

$$\Delta\eta_{ij}=\eta_{ij}(E)-\eta_{ij}(0)=\sum_k\gamma_{ijk}E_k. \tag{5}$$

线性电光效应不可能出现在具有空间反演的晶格结构中. 反演操作 I 使所有的矢量变号. 对(5)式施行反演操作,得到

$$\Delta\eta'_{ji}=I\Delta\eta_{ij}=\sum_k\gamma_{ijk}(-E_k)=-\sum_k\gamma_{ijk}E_k, \tag{6}$$

由于反演对称性,所以有

$$\Delta\eta'_{ij}=\Delta\eta_{ij}, \tag{7}$$

(6)式和(7)式同时成立的条件是所有的 $\gamma=0$. 因此在11种具有反演对称中心的晶族中所有的三阶张量 $\gamma_{ijk}=0$,只有那些不存在反演对称中心的晶体中才有可能出现线性电光效应.

借助于缩写符号,我们可以把法线椭球方程式(2)写成

$$\begin{aligned}\sum_{ij}\eta_{ij}(E)\xi_i\xi_j=\sum_k\Big\{&\left(\frac{1}{n_x^2}+\gamma_{1k}E_k\right)x^2+\left(\frac{1}{n_y^2}+\gamma_{2k}E_k\right)y^2\\&+\left(\frac{1}{n_z^2}+\gamma_{3k}E_k\right)z^2+2\gamma_{4k}E_kyz\\&+2\gamma_{5k}E_kzx+2\gamma_{6k}E_kxy\Big\}=1,\end{aligned} \tag{8}$$

式中 $n_i(i=x,y,z,)$ 为未加电场时的主折射率,xyz 为未加电场时的晶体主轴坐标系. 当外场为0时,上式化为普通的法线椭球方程式

$$\frac{x^2}{n_x^2}+\frac{y^2}{n_y^2}+\frac{z^2}{n_z^2}=1. \tag{9}$$

在外电场的作用下,法线椭球的主轴一般并不与原来的晶体主轴 x,y,z 重合. 通过正交变换,可将(8)式化为标准型,从而确定外电场感生的主轴方向.

晶体点群的对称操作使 6×3 阶张量 $\gamma_{\mu i}$ 的独立元素进一步减少. 例如,KDP 晶体(KH_2PO_4,磷酸二氢钾)为单轴晶体,属于四方晶系的 $\bar{4}2\,m$ 点群,取它的四重轴为 z 轴(即光轴),它的电光系数张量为

$$\gamma = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \gamma_{41} & 0 & 0 \\ 0 & \gamma_{41} & 0 \\ 0 & 0 & \gamma_{63} \end{pmatrix}, \tag{10}$$

仅有的三个非零元素为 $\gamma_{41} = \gamma_{52}$ 及 γ_{63}，加电场后变成双轴晶体，以 γ_{41}，γ_{63} 代入(8) 式得到

$$\frac{x^2}{n_o^2} + \frac{y^2}{n_o^2} + \frac{z^2}{n_e^2} + 2\gamma_{41}E_x yz + 2\gamma_{41}E_y zx + 2\gamma_{63}E_z xy = 1. \tag{11}$$

如果外电场沿光轴方向，上式变成

$$\frac{x^2 + y^2}{n_o^2} + \frac{z^2}{n_e^2} + 2\gamma_{63}E_z xy = 1. \tag{12}$$

方程(12) 可通过正交变换(将坐标系统 z 轴旋转 45°)

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \xi \\ \eta \end{pmatrix}, \tag{13}$$

化为标准型

$$\left(\frac{1}{n_o^2} + \gamma_{63}E_z\right)\xi^2 + \left(\frac{1}{n_o^2} - \gamma_{63}E_z\right)\eta^2 + \frac{\zeta^2}{n_e^2} = 1, \tag{14}$$

式中 $\zeta = z$. 上式可写作

$$\frac{\xi^2}{n_\xi^2} + \frac{\eta^2}{n_\eta^2} + \frac{\zeta^2}{n_\zeta^2} = 1. \tag{15}$$

式中 n_ξ，n_η 和 n_ζ 是新的主轴坐标系下的折射率，它们精确到一级小量的表达式为

$$\left.\begin{aligned} n_\xi &= n_o - n_o^3\gamma_{63}E_z/2, \\ n_\eta &= n_o + n_o^3\gamma_{63}E_z/2, \\ n_\zeta &= n_e. \end{aligned}\right\} \tag{16}$$

这样外加电场的方式称"纵调制"，表明外电场与晶体原来的光轴方向一致.

砷化镓(GaAs)属于$\bar{4}3m$点群，锗酸铋(BGO，$Bi_{12}GeO_{20}$)和硅酸铋(BSO，$Bi_{12}SiO_{20}$)属于23点群，都是立方晶系，未加电场时为光学各向同性介质. 加电场后变为双轴晶体，它们的电光系数张量为

$$\gamma=\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\\ \gamma_{41} & 0 & 0\\ 0 & \gamma_{41} & 0\\ 0 & 0 & \gamma_{41}\end{pmatrix}, \tag{17}$$

代入(8)式得到

$$\frac{x^2+y^2+z^2}{n^2}+2\gamma_{41}E_z xy=1. \tag{18}$$

运用正交变换(13)化为标准型

$$\left(\frac{1}{n^2}+\gamma_{41}E_z\right)\xi^2+\left(\frac{1}{n^2}-\gamma_{41}E_z\right)\eta^2+\frac{1}{n^2}\zeta^2=1, \tag{19}$$

在电场感生的主坐标系$\xi\eta\zeta$下三个主折射率为

$$\left.\begin{aligned}n_\xi&=n-\frac{1}{2}n^3\gamma_{41}E_z,\\ n_\eta&=n+\frac{1}{2}n^3\gamma_{41}E_z,\\ n_\zeta&=n.\end{aligned}\right\} \tag{20}$$

铌酸锂($LiNbO_3$)、钽酸锂($LiTaO_3$)的点阵对称性为$3m$，它们均有一个三重对称轴，即光轴(z轴)，其电光系数张量为

$$\gamma=\begin{pmatrix}0 & -\gamma_{22} & \gamma_{13}\\ 0 & \gamma_{22} & \gamma_{13}\\ 0 & 0 & \gamma_{33}\\ 0 & \gamma_{51} & 0\\ \gamma_{51} & 0 & 0\\ -\gamma_{22} & 0 & 0\end{pmatrix}, \tag{21}$$

三个主折射率为

$$\left.\begin{aligned} n_\xi &= n_o - \frac{1}{2}n_o^3\gamma_{13}E_z, \\ n_\eta &= n_o - \frac{1}{2}n_o^3\gamma_{13}E_z, \\ n_\zeta &= n_e - \frac{1}{2}n_e^3\gamma_{33}E_z. \end{aligned}\right\} \tag{22}$$

7.6.2 电光晶体的本征模和合成偏振态

我们已讨论了外电场引起的晶体法线椭球的变化. 下面我们以 BSO 为例,进一步研究平面偏振光在电光介质中的传播.

BSO 原来是光学各向同性介质,加电场后变成双轴晶体,过原点且与电场方向(ζ 向)垂直的平面与椭球相交成椭圆. 其半轴方向即ξ与η方向. 这样一来,简并退化,沿ζ轴只能传播一对正交的本征模,分别在ξ,η方向偏振,折射率由(20)式表示. 当光波在ζ方向传播的距离为L时,两个本征模的相位差

$$\Gamma = \frac{2\pi}{\lambda}(n_\eta - n_\xi)L = \frac{2\pi}{\lambda}n^3\gamma_{41}E_zL = \frac{2\pi}{\lambda}n^3\gamma_{41}V, \tag{23}$$

式中 $V = E_zL$ 为外加电压值.

当入射光波在x方向(即ξ,η的角平分线)方向偏振时,在$\xi\eta$坐标系内的 Jones 矩阵[6] 为

$$J = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \tag{24}$$

射出晶体后 Jones 矩阵变为

$$J' = \frac{1}{\sqrt{2}}\begin{bmatrix}e^{i\Gamma/2}\\e^{-i\Gamma/2}\end{bmatrix}, \tag{25}$$

两个本征模在晶体中获得的相位差Γ由(23)式表示. 一般情况下合成偏振态为椭圆偏振光. 通常把$\Gamma = \pi$时的外加电压称为半波电压,记为V_π,由(23)式,

$$V_\pi = \lambda/2n^3\gamma_{41}. \tag{26}$$

7.6.3 振幅调制和相位调制

设晶体是与 xy 平行的晶片，沿 z 方向的厚度为 L，在输出端放一个与 x 方向平行的起偏振器，入射光波沿 z 方向传播，且沿 x 方向偏振，射入晶体后，它分解成 ξ,η 方向的偏振光(图 7.19)，射出晶体后的偏振态由 (25) 式表示. 如果在输出端放一个与 y 平行的检偏振器，就构成 Pockels 盒. 则由检偏器输出的光波 Jones 矩阵为



图 7.19 xy 坐标系和 $\xi\eta$ 坐标系(感生坐标系)

$$J''=\frac{1}{\sqrt{2}}\begin{pmatrix}-1 & 1\\ 0 & 0\end{pmatrix}J'=\frac{1}{2}\begin{pmatrix}-e^{i\Gamma/2}+e^{-i\Gamma/2}\\ 0\end{pmatrix}, \tag{27}$$

它表示输出光波是沿 y 方向的线偏振光，其光强为

$$I''=\frac{I_0}{2}(1-\cos\Gamma)=I_0\sin^2\left(\frac{\pi V}{2V_\pi}\right). \tag{28}$$

上式说明光强受到外加电压的调制，称振幅调制，I_0 为光强的幅值，当 $V=V_\pi$ 时 $I''=I_0$. 图 7.20 为振幅调制示意图，Pockels 盒的输出随着外电压的加大而加大，表明有更多的能量从 x 偏振态转移到 y 偏振态中去.

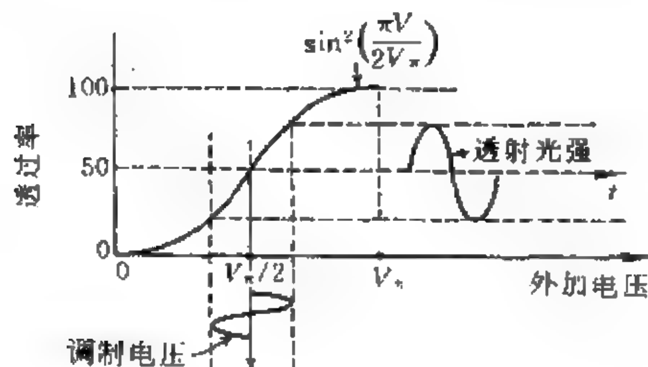


图 7.20 线性电光效应振幅调制器的特性曲线[3]

如果入射光波预先加四分之一波片移相，则

$$I'' = \frac{I_0}{2}\left[1 - \cos\left(\Gamma + \frac{\pi}{2}\right)\right] = I_0\sin^2\left(\frac{\pi V}{2V_\pi} + \frac{\pi}{4}\right), \quad (29)$$

加上预置的相位后，工作点移到调制曲线的中点附近，使线性大大改善.

如果输入光波的振动平面沿 ξ 方向，则沿 z 轴方向的外电场 V 仅仅使光波产生一个附加相位 $\Delta\varphi$. 由(10)式，在调制晶体为 BSO 的情况下，

$$\Delta\varphi = \frac{2\pi}{\lambda}n_\xi L = \frac{2\pi}{\lambda}nL - \frac{\pi}{\lambda}n^3\gamma_{41}V = \frac{2\pi}{\lambda}nL - \frac{\pi V}{2V_\pi}. \quad (30)$$

外加电压调制了附加相位的大小，称相位调制.

7.6.4 PROM 器件的结构及工作原理

PROM 器件是 Pockels readout optical modulator 的简写，其结构如图 7.21 所示. 它利用 BSO 晶体作为调制介质，其表面垂直于(001)方向，厚度约数百 μm. 它的两边分别有一个绝缘层 INS_1，INS_2，通常是对二甲苯基，厚度约几个 μm；在 BSO 和 INS_2 之间有一个二向色反射镜，镀上短波通膜层，它能反射红光、透射蓝光；绝缘层外面用溅射法涂以氧气铟锡，构成透明电极. BSO 晶体既是电光介质，又是光电导介质，在光波的照射下，容易产生电子-空穴对. 图 7.21 中 I_R，I'_R 为入射和反射的红光；I_B 为入射的蓝光. 图中 ITO_1，ITO_2 为透明电极；INS_1，INS_2 为绝缘层；M 为二向色反射镜(反射红光，透过蓝光)；BSO 为调制晶体.

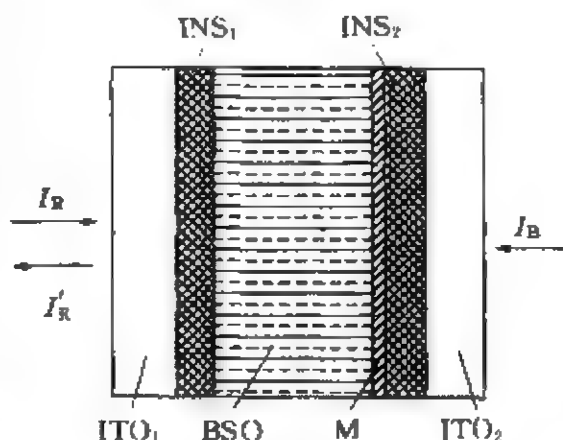


图 7.21 PROM 器件构造示意图

器件首先加上直流电压 V_0，其数值约为 V_x 的一半，这一电压被加到 BSO 晶体及两个绝缘层上，如图 7.22(a)所示．如果晶体被脉冲氙灯均匀照射，则由于 BSO 的光电导效应，其中产生大量电子-空穴对，使 BSO 导电，从而 BSO 内的压降为 0，如(b)所示．关掉氙灯后，突然将 PROM 两测电源短路，则晶体上将出现 $-V_0$ 的压降．见图 7.22(c)．在这种情况下，我们再用短波长的蓝色图像从右边照射晶体(参见图 7.24)．蓝色图像将透过二向色反射

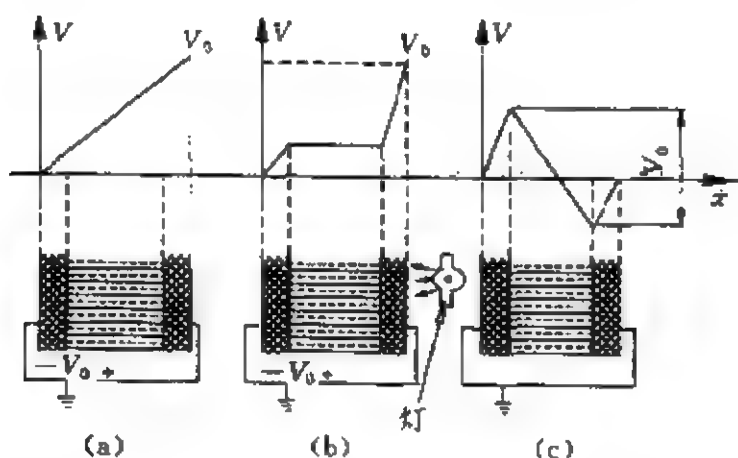


图 7.22　PROM 的工作原理

镜，射入 BSO 晶体，在图像的亮区，由于光电导效应产生电子-空穴对，电子在电场的作用下向 BSO 的左端面移动，正、负电荷分离后形成的内电场将抵消一部分外电场使这些区域的压降减小；在暗区则由于电子-空穴对很少，压降基本维持 V_0 不变．这样一来，空间光强分布经光电导效应转换成空间的电压分布，或者说用蓝色光写入的光强图样转化为电压图样．电压 V 与曝光量 E 的关系是指数关系而不是线性关系，即

$$V=V_0e^{-KE}. \tag{31}$$

图 7.23 描绘了这一关系，即 BSO 上暗区的压降大，亮区的压降小．(31)式中 K 为常数．

读出光一般用长波长的红光，例如 He-He 光．光电导效应的

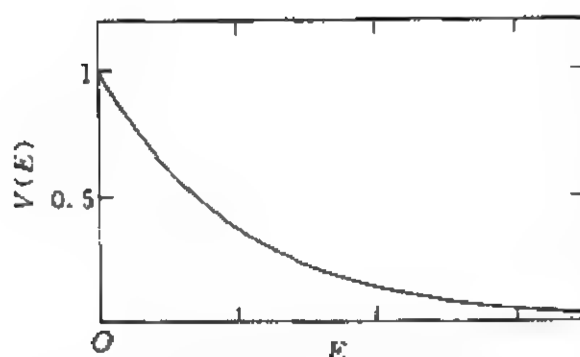


图 7.23　PROM 器件的电压降与曝光量的关系

灵敏度强烈地依赖于光波波长，对于 BSO 晶体，$\lambda=633$ nm 的红光的灵敏度仅为 $\lambda=400$ nm 蓝紫光灵敏度的 1/200 左右，因此可以近似认为读出光不产生光电导效应，不会影响蓝光写入的电压图形.

He-Ne 激光为线偏振光，振动方向沿 x 轴，它透过 BSO 后被二向色反射镜 M 反射，再透过 BSO 晶体，由于 Pockels 效应，在不同的区域，由于不同的电压降 V，引起了不同的相位差，如果在输出端加上沿 y 方向的检偏器，则输出光强将遵循公式(28)，再度将电压分布转换成光强分布，或者说红光读出了原来蓝光写入的图形. 这样就实现了光寻址的空间光调制器的功能. PROM 器件是通过短波长光的光电导效应写入图像(把空间光强分布转换成空间电压分布)，又通过长波长光的线性电光效应读出图像(把空间电压分布恢复成空间光强分布). 对于理想的 PROM 器件，在写入图像的暗区，读出光两次经过 BSO 获得最大的光强透过率；在亮区，读出光的透过率很低. 因此读出的是负像.

PROM 器件对晶体均匀性的要求很高. 此外，尽管读出光红光引起的光电导效应并不显著，但它毕竟使晶体上的电压有所降低，因而 PROM 器件不能在较强的读出光照明下长时间工作.

图 7.24 为 PROM 器件的读出光路，从 He-Ne 激光器辐射的线偏振光束与偏振分光棱镜的主截面平行，为 p 分量，它通过偏振分光棱镜 PBC 后射入 PROM 器件，振动方向沿 BSO 晶体的 x 方

向，在晶体中往返一次后，电压图像转移到光波中成为二维的光强分布，然后再由PBC反射. 在图像的暗区，光波获得的相位差$\Gamma\approx\pi$，反射光成为y方向的线偏振光，即PBC的s分量，在PBC中接近完全反射，而在亮区晶体的电压却被光生电荷产生的电场所抵消，$\Gamma<\pi$. 由(28)式，PBC对它的反射率较低. 因此由PBC反射的正是蓝光写入图像的负像. 通常写入光是非相干光I_w，而读出光是相干光，因此PROM器件可实现非相干光-相干光的图像转换.

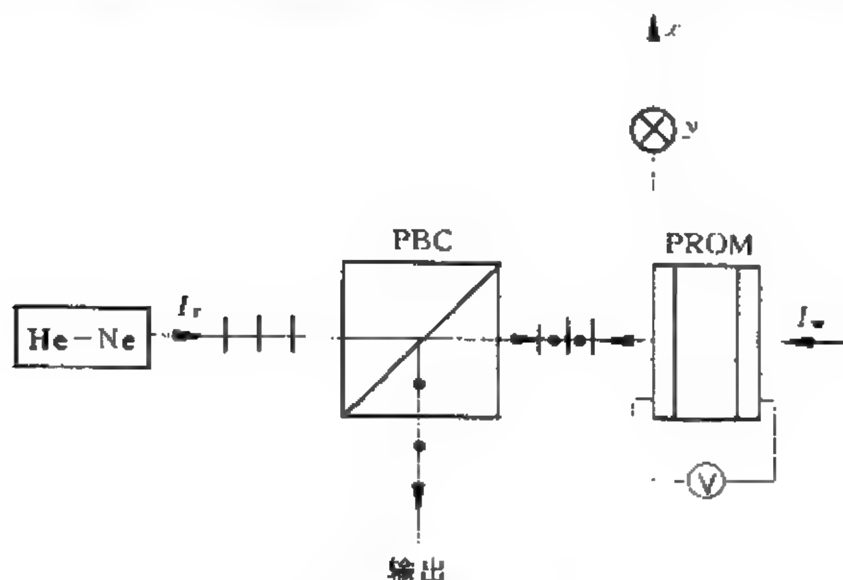


图 7.24　PROM器件的读出光路

PROM器件还用BGO，ZnSe作为调制晶体，与BSO类似，它们既是线性电光介质又是光电导介质. 但这两种晶体的半波电压较高($V_{\pi,BGO}=5.5$ kV，$V_{\pi,ZnSe}=11.3$ kV)，而BSO的半波电压较低(3.9 kV)，因此近年来PROM器件主要用BSO制造.

7.7　数字微反射镜器件(DMD)和数字化投影

近年来，一种利用微细加工和大规模集成电路技术、构思极为巧妙的电寻址空间光调制器——数字微反射镜器件(digital micromirror device)问世了. 这种器件原来称作变形反射镜器件(de-

formable mirror device)，两种名称的英文缩写恰好都是 DMD，它是由美国德克萨斯仪器公司(TI)的一名科学家 L. J. Hornbeck 在 1987 年发明的，DMD 作为光学信息处理系统接口器件的报道并不多，但是近年来，它应用于数字投影显示(digital projection display)及高清晰度电视(HDTV)中显示出来的优越性能，却引起了科技界和工业界的广泛兴趣.

7.7.1 DMD 的结构和工作原理

DMD 的结构如图 7.25 所示. 器件的基底是硅，用大规模集成电路的技术，在硅片上制出 RAM，每一个存储器有两条寻址电极(addressing electrodes)和两个搭接电极(landing electrodes). 两个支撑柱上，通过扭臂梁铰链(torsion hinge)安装一个微形反射镜，形成一个“跷跷板”的结构.

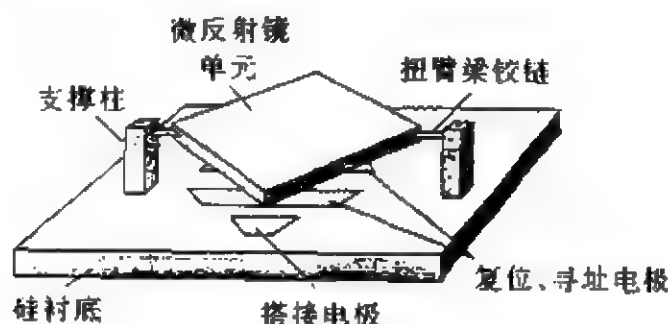


图 7.25 DMD 的结构示意图[19]

器件工作时，在反射镜上加负偏压，一个寻址电极上加+5 V(数字“1”)，另一寻址电极接地(数字“0”)，这样一来，就形成一个差动电压，它产生一个力矩，使反光镜绕扭臂梁旋转，直到触及搭接电极为止. 在扭转力矩的作用下，反射镜将一直锁定于这一位置，不管它下面的存储器的数据是否变化，直到复位信号出现为止，对应旋转角 $\theta_L = 10°$. 这样一来，每一单元都有三个稳态：+10°、-10°和 0°，$\theta = 0°$对应于没有寻址信号(两个寻址电极都是 0)的情况. DMD 是通过半导体微细加工技术精密制作的，因此反

射镜列阵的三个稳态一致性相当好，对应于 DMD 的三个平面：与基平面成$\pm\theta_L$角的倾斜平面及平行于基面的平面.

作为非相干光调制器，DMD 用于投影仪时工作原理参见图 7.26. 光源发出的光束与光学系统光轴的夹角为$2\theta_L$，倾斜照射

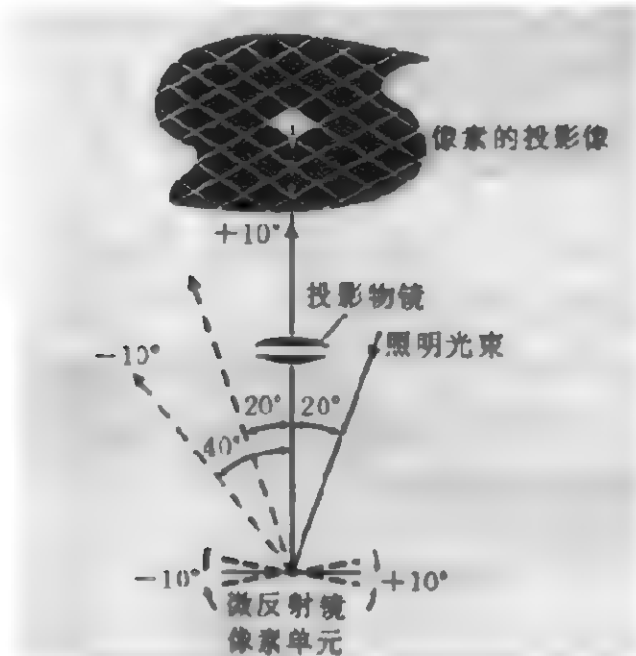


图 7.26 DMD 作为数字投影仪空间光调制的工作原理图[19]

DMD，当某一像素的反射镜$\theta=0$或$\theta=-\theta_L$时，反射光通不过投影物镜. 当该像素被寻址时，$\theta=\theta_L$，它反射的光束正好沿光轴方向通过投影物镜，成像到屏上，称此状态为"开(ON)"；$\theta=-\theta_L$则对应于 DMD 的状态"关(OFF)". DMD 一般由视频信号驱动. 在每一帧的时间内，某一像素处于两种状态的占空比，决定了该像素的灰阶，亦即灰阶是由入射光的二元脉冲宽度调制实现，人的眼睛再把这一"数字"图像翻译成"模拟"图像. 一般灰阶数为$2^8=256$. 颜色则通过两种方式加到图像中去，一个方法是在照明光路中加一个 R,G,B 三原色滤色镜的调色盘，它与视频信号严格同步，在每一帧的时间Δt内转一圈，在各色盘范围内再分别用像素 ON/OFF 的占空比调节 R,G,B 的强度比，从而在一帧的时间Δt内合成所要求的颜色. 由于每种颜色的灰阶都是 8 bit 即 256 种，总共

可产生 256^3 约 1600 万种不同的颜色，色彩是相当丰富的.

7.7.2 DMD 制作工艺简介

DMD 用大规模集成电路技术制作，大体上可分为以下步骤：

(1)用标准的 CMOS 工艺在硅片上制作记忆单元、寻址电极等.

(2)在硅片表面覆盖一层高分子聚合物，其厚度相当于反射镜的高度.

(3)用光刻、溅射法刻透聚合物层，制作支撑柱.

(4)镀一层薄的铝膜，作为扭臂梁层，镀一层厚的铝膜，作为反射镜层；用光刻法将扭臂梁和反射镜成形.

(5)用离子刻蚀法除去余下的高分子聚合物层，最后形成架空的微反射镜.

当出现寻址信号及负偏压时，薄的扭臂梁扭转变形，厚的反射镜则不变形，只作整体的偏转.

7.7.3 DMD 数字式投影仪

作为空间光调制器，DMD 可应用于相干光信息处理[21]，然而它更重要的用途是作为非相干光的空间光调制器. 例如用在打印机上[20]，特别是用于投影系统，产生了与传统的投影仪完全不同的数字式投影仪.

图 7.27 是用 DMD 作为空间光调制器的 110 英寸大屏幕数字式投影仪的系统图，它事实上也是一台大屏幕投影电视机. 由光源发出的白光宽光束经过第一个聚光镜照射三原色调色盘，它在视频信号的控制下同步旋转，在每一帧的时间内旋转一圈，经调色盘调制的光束照亮 DMD 器件. 从器件上反射的光束经变焦物镜 ZOOM 投射到屏幕上，屏幕上投影像各点的灰度及颜色是由对应的微型反射镜处于开态(ON)的时间决定的，而开态的时间长短则可以用脉冲宽度加以调制. 通常通过 8 bit 的系统为每种颜色提

供 256 级灰度，每种颜色对应的总时间占一帧的 1/3，因此色彩的总数为 256^3. 由于视觉暂留效应，人眼所看到的是三原色合成的彩色图像，这一工作方式称为“顺序颜色模式”(sequential color mode).

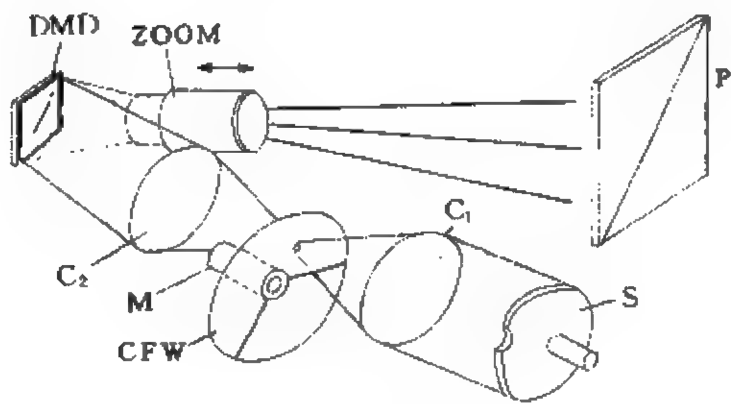


图 7.27 DMD 数字式投影仪[19]

C_1,C_2 为聚光镜；M 为电机；CFW 为三原色滤光镜盘

P 为 110 英寸屏幕；S 为光源

更先进的系统是所谓三板式 DMD 投影机. 参见图 7.28. 由短弧、高强度、高色温光源(金属化卤钨灯或氙灯)发出的白光经反射镜、聚光镜及折转反射镜后，通过一个颇为复杂的分光棱镜系统分成 R,G,B 三色光，分别以 20°角(即 $2\theta_L$ 角)照射三个 DMD，三个 DMD 用三色电信号分别驱动，三个 DMD 器件中处于 ON 态的像素的反射光再通过该棱镜系统重新合成，通过一个变焦物镜

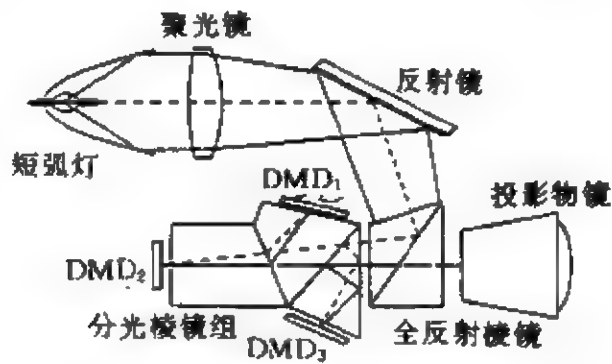


图 7.28 3-DMD 式投影系统[22]

投影到屏幕上．由于 3-DMD 投影系统中各单色光分别用各自的 DMD 调制，因而其信号是连续的，亮度更高(理论上是单 DMD 系统的 3 倍)，而灰阶更丰富，可达 2^{11}(11 bit)，因此可以合成更多种类的颜色．3-DMD 系统适用于超大屏幕或需要更明亮的投影像的情形，但三个 DMD 的空间位置必须仔细调整，使它们的图像精确重合．

7.7.4 DMD 数字投影技术的特点

DMD 空间光调制器的原理，也许是我们所介绍的各种空间光调制器中最简单的．效应越简单，一般来说附属设备及系统结构就越紧凑，其结果，是高效率及高度的稳定性．此外，DMD 可以用成熟的大规模集成电路技术制造，这就为它的商品化创造了条件．巧妙的构思加上集成电路的制造工艺，使 DMD 在分辨率、亮度、反差、灰阶、色保真度及响应时间等主要指标上都达到了目前显示技术的创记录的水平．

表 7.3 给出用 DMD 作空间光调制器的单板和三板数字投影仪的主要性能指标．DMD 数字投影技术具有以下特点：

(1) 高分辨率

这是由 DMD 的像素尺寸及制造工艺决定的．DMD 可以有 640×480(VGA)、800×600(SVGA)、1024×768(XGA)及 1280×1024(SXGA)等多种解析度规格，特别是适用于 16：9 宽屏幕电视的 DMD，器件尺寸 37 mm×22 mm，像素尺寸为 17 μm×17 μm，一个 DMD 上的像素数为 2048×1152，远远超过 LCD 的像素数，达到 N 制电视制式器件的五倍以上，完全符合高清晰度电视(high definition TV，HDTV)的要求．

(2) 高亮度

DMD 不存在 LCD 器件所必需的偏振片组，它完全靠反射镜阵列的反射．反射镜面积约为像素面积的 70%，两个像素的间隔为 1 μm，再扣除光在扭臂梁、支撑柱及反射镜边缘的散射、衍射，有效反射率达到 61%，远远高于 LCD 的光学效率．

表 7.3 应用 DMD 的数字投影系统[19]

指 标	单板(1-DMD)投影仪 NTSC 标准制式显示	三板(3-DMD)投影仪高清晰度电视(HDTV)
DMD 的像素数	768×576	2048×1152
有效像素数	640×480	1707×910
像素尺寸	16 μm×16 μm	16 μm×16 μm
像素周期	17 μm	17 μm
像素有效孔径*	74%	>75%
投影方式	前向投影	前向投影
屏幕尺寸(对角线)	≥52 英寸	≥60 英寸
画面长宽比	4∶3	16∶9
颜色帧频	180 Hz	60 Hz
屏幕亮度(光通量)	220 lm	>2000 lm
光源	1000 W，氙灯	575 W，金属化卤钨灯
对比度	>50∶1	>100∶1
每种颜色灰阶	8 bit(256 级)	8 或 9 bit
颜色生成方式	同步旋转调色盘	二向色棱镜组
像素对准	自对准	光学对准

* 即开口率，反射镜面积与像素总面积之比，又称 fill factor.

(3) 对比度、灰阶及色保真度

在 DMD 的关态(OFF)下从反射镜上反射的光束与光轴的夹角为 $4\theta_L$，约 40°，被仪器四壁专门设计的吸收介质吸收，只有少量散射光进入投影物镜，因此对比度达到 50∶1 是没有问题的，目前改良的技术已达到 100∶1 以上.

作为一个数字化器件，它的灰阶及色保真度由开态(ON)的占空比决定，而占空比由脉冲调制来控制，一般能做到 8 bit(256 级)甚至 10 bit(1024 级)，上面已谈过，三原色各 8 bit 的混合结果产生 256^3 即 1600 万种不同的颜色，无论是灰阶还是色保真度都达到 HDTV 的要求.

7.3 节图 7.7 给出液晶器件的电光特性曲线，它表明 LCD(或 LCLV)是一个模拟式空间光调制器，而 DMD 则是数字式空间光调制器. 数字化是今后信息技术发展的大势所趋. 也是电视

技术发展的必然趋势. 因此在大屏幕投影技术的竞争中，DMD作为数字式空间光调制器件，无疑占了有利的地位.

(4) 可靠性

由于 DMD 的调制功能是通过微型反射镜的扭转实现的，因此人们自然会担心在反复的转动下扭臂梁铰链的可靠性. TI 公司专门做了模拟 DMD 长期使用的试验，历经 765×10^4 个周期未发生任何问题，这相当于投影电视或投影仪正常使用 76 000 小时，看来可靠性是不成问题的.

(5) 响应时间

由于扭臂梁非常薄(约 0.05～0.1 μm)，微型反射镜的重量很轻，转动惯量极小，因此响应时间极短，从完全的开态到完全的关态约 10 μs，因此单板式 DMD 投影仪可以在常规 TV 的 1/3 帧的时间内读完全部 RGB 三色信号.

第一台 DMD 投影系统已于 1996 年投放市场，专业级或消费级的 DMD 投影电视正在进入商品化阶段.

在本世纪末和即将来临的 21 世纪初，高清晰度电视(HDTV)将逐步取代常规的电视，由此，空间光调制器必须具有 200 万以上的像素，高达 30 MHz 以上的数据率(data rate). HDTV 和大屏幕显示常常是同时要求的. 此外，数字化也是 HDTV 的发展趋势. DMD 几乎满足 HDTV 的所有指标，可以预言，DMD 是未来 HDTV 首选的空间光调制器件. 因而有人把 DMD 称为"数字化投影技术的革命".

有关 DMD 的详细介绍，请参看 Younse[19]，Florence[22] 和 Hornberck[23]的论文.

参 考 文 献

[1] Y. Ichioka, T. Iwaki and K. Matsuoka, Optical Information Processing and Beyond, Proc, *IEEE. J. Quantum Electron.*, 84(1996), 694-719.

[2] F. T. S. Yu and S. Jutamulia, Optical Signal Processing, Computing, and Neural Networks, John Wiley & Sons, New York, 1992.

[3] P. Yeh, Optical Waves in Crystals, John Wiley & Sons, New York, 1984.

[4] 李育林,傅晓理,空间光调制器及其应用,国防工业出版社,1996.

[5] 赵达尊,张怀玉,空间光调制器,北京理工大学出版社,1992.

[6] 羊国光,宋菲君,高等物理光学,中国科技大学出版社,1991.

[7] S. E. Shields, Optics & Photonics News., **5**(1994),12.

[8] J. Grinberg et al, A new real-time noncoherent to corerentlight image converter, the hybrid field effect liquid crystal valve, *Opt. Eng.*, **14**(1975), 217-225.

[9] W. P. Bleha et al, Application of the liquid crystal light valve to real-time optical data processing, *Opt. eng.*, **17**(1978), 371-384.

[10] 黄锡珉,平板显示专题系列介绍——液晶显示,物理,**20**(199),366-370.

[11] 李清泉,彩色电视液晶显示器的发展,物理,**20**(199),114-117.

[12] 周连祥,平板显示器评论,物理,**22**(199), 301-305.

[13] U. Efron et al., Silicon liquid crystal light valves: status and issues, *Opt. Eng.*, **22**(1983), 682-686.

[14] U. Efron et al., the silicon liquid crystal light valves, *J. Appl. Phys.*, **57**(1985), 1356-1368.

[15] S. Fukushima and T. Kurokawa, Ferroelectric liquid crystal spatial light modulators and their applications to information processing, Tech. Dig. OSA Topical Meet. on Spatial Light Modulators and Applicat., 1995, 130-134.

[16] D. Amitage, J. I. Tackara and W. D. Eades, Photoaddresed spatial liquid crystal light modulators, *Appl. Opt.*, **28**(1989), 4763-4774.

[17] D. Willians, An amorphous silicon/chiral smectic spatial light modulator, *J. Phys. D.*, **21**(1988), 5156-5159.

[18] G. Moddel et al., Hight speed binary optical addressing spatial light modulators, *Appl. Phys. Lett.*, **55**(1989), 537-539.

[19] J. M. Younse, Mirrors on a chip, *IEEE Spectrum*, **11**(1993), 27-31.

[20] E. Nelson and L. Hornbeck, Micromechanical spatial light modulator for electrophotographic printers, in Proc. SPSE, fourth International Congress in Non-Impact Printing technologies, (1988), 427.

[21] J. M. Florence and R. Gale, Coherent optical correlation using a deformable mirror device spatial light modulator in Fourier plane, *Appl. Opt.*, **27**(1988), 2091-2092.

[22] J. M. F.orence and L. A. Yoder, Display system architectures for digital micromiror divice(DMD) based projectors, Proc., SPIE 2650, (1996)193-208.

[23] L. J. Hornberck, Digital light processing for high-brightness, high-resoklution applications, *Proc.*, *SPIE*3013, *(1997)*, *27-41*.

[24] S. Jutamulia and F. T. S. Yu, SLMs act as an instant phtographic transparency, OE Reports 161, May(1997).

[25] P. R. Barbier and G. Moddel, Spatial light modulators: processing light in real time, *Opt. Photonics News*, **8**(1997), 17-21.

附录 A　偏振光在线性扭曲介质中的传播

设线性扭曲介质(液晶)被限制在平面 $z=0$ 和 $z=d$ 之间,介质的光轴与 xy 平面平行,并随着 z 线性变化,其方位角

$$\psi=\alpha z, \tag{1}$$

α 为常数. 经过厚度为 d 的介质后,

$$\psi=\psi_d=\alpha d. \tag{2}$$

设介质表面($z=0$)处光轴的方向取为 x 方向,把介质均匀分成 N 层,均与 xy 平面平行,N 足够大,每一层的厚度 Δz 足够小,从而每一层中的局部光轴可以认为是不变的. 光波沿 z 轴入射,在 $z=0$ 处光波的振动平面沿 x 方向(光轴方向).

第 n 层的方位角

$$\psi_n=n\Delta\psi=n\psi_d/N, \tag{3}$$

其中 $\Delta\psi$ 是每一层中光轴的旋转角,$\Delta\psi=\psi_d/N$.

设入射到第 n 层的光波偏振的 Jones 矩阵[6]为 $\begin{pmatrix}E_x\\E_y\end{pmatrix}$,它将分解成该层的两个主方向(快轴 f 与慢轴 s)的分量:

$$\begin{pmatrix}E_s\\E_f\end{pmatrix}=R(\psi_n)\begin{pmatrix}E_x\\E_y\end{pmatrix}, \tag{4}$$

其中

$$R(\psi_n)=\begin{pmatrix}\cos\psi_n & \sin\psi_n\\ -\sin\psi_n & \cos\psi_n\end{pmatrix}, \tag{5}$$

表征旋转 ψ_n 角的矩阵. 这一层介质对光波的作用相当于一个相位延迟片,经过它的作用,光波的偏振态为

$$\begin{aligned}\begin{pmatrix}E_s'\\E_f'\end{pmatrix}&=\begin{pmatrix}\exp(\mathrm{i}2\pi n_s\Delta z/\lambda) & 0\\ 0 & \exp(\mathrm{i}2\pi n_f\Delta z/\lambda)\end{pmatrix}\begin{pmatrix}E_s\\E_f\end{pmatrix}\\&=\exp(\mathrm{i}\Gamma'/2N)W_0\begin{pmatrix}E_s\\E_f\end{pmatrix},\end{aligned} \tag{6}$$

式中 $n_s=n_o$，$n_f=n_e$，且有

$$\Gamma'=2\pi(n_e+n_o)d/\lambda, \tag{7}$$

$$W_0=\begin{pmatrix}\exp(-\mathrm{i}\Gamma/2N) & 0\\ 0 & \exp(\mathrm{i}\Gamma/2N)\end{pmatrix}, \tag{8}$$

$$\Gamma=2\pi(n_e-n_o)d/\lambda. \tag{9}$$

将(6)式代回固定坐标系，得到

$$\begin{aligned}\begin{pmatrix}E'_x\\ E'_y\end{pmatrix}&=R(-\psi_n)\begin{pmatrix}E'_s\\ E'_f\end{pmatrix}\\ &=\exp(\mathrm{i}\Gamma'/2N)R(-\psi_n)W_0R(\psi_n)\begin{pmatrix}E_x\\ E_y\end{pmatrix}\\ &=\exp(\mathrm{i}\Gamma'/2N)W_n\begin{pmatrix}E_x\\ E_y\end{pmatrix},\end{aligned} \tag{10}$$

式中

$$W_n=R(-\psi_n)W_0R(\psi_n). \tag{11}$$

介质的总体效应可表为

$$\begin{aligned}W&=\exp(\mathrm{i}\Gamma'/2)\prod_{n=1}^{N}W_n\\ &=\exp(\mathrm{i}\Gamma'/2)R(-N\Delta\psi)W_0R(N\Delta\psi)\\ &\quad\times R[-(N-1)\Delta\psi]W_0R[(N-1)\Delta\psi]\\ &\quad\times\cdots\times R(-\Delta\psi)W_0R(\Delta\psi)\\ &=\exp(\mathrm{i}\Gamma'/2)R(-\psi_n)[W_0R(\Delta\psi)]^N,\end{aligned} \tag{12}$$

其中用到 $R(-N\Delta\psi)=R(-\psi_d)$ 及 $R(\alpha)R(\beta)=R(\alpha+\beta)$. 记

$$W_0R(\Delta\psi)=\begin{pmatrix}W_{11} & W_{12}\\ W_{13} & W_{14}\end{pmatrix}, \tag{13}$$

其中

$$\left.\begin{aligned}W_{11}&=\cos(\varepsilon\psi_d)\mathrm{e}^{-\mathrm{i}\varepsilon\Gamma/2},\\ W_{12}&=\sin(\varepsilon\psi_d)\mathrm{e}^{\mathrm{i}\varepsilon\Gamma/2},\\ W_{21}&=-\sin(\varepsilon\psi_d)\mathrm{e}^{-\mathrm{i}\varepsilon\Gamma/2},\\ W_{22}&=\cos(\varepsilon\psi_d)\mathrm{e}^{\mathrm{i}\varepsilon\Gamma/2},\end{aligned}\right\} \tag{14}$$

式中 $\varepsilon=1/N$，当 $N\to\infty$时，$\varepsilon\to 0$，从而 W_{11}和 W_{12}可近似表为

$$W_{12}\approx -W_{21}\approx \sin\varepsilon\psi_d. \tag{15}$$

这样一来，我们就可以应用 Chebyshev 等式(参见(21)式)，得到

$$[W_0R(\Delta\psi)]^N = \frac{1}{\sin\eta}\begin{bmatrix} W_{11}\sin N\eta-\sin(N-1)\eta & W_{12}\sin N\eta \\ W_{21}\sin N\eta & W_{21}\sin N\eta-\sin(N-1)\eta \end{bmatrix}, \tag{16}$$

其中

$$\eta=\arccos[(W_{11}+W_{22})/2], \tag{17}$$

且有下面的近似公式

$$\left.\begin{aligned} &\cos\eta\approx 1-\varepsilon^2X^2/2, \\ &X^2=\psi_d^2+(\Gamma/2)^2, \end{aligned}\right\} \tag{18}$$

以及

$$\left.\begin{aligned} &\sin\eta\approx\eta\approx\varepsilon X, \\ &\sin N\eta\approx\sin X. \end{aligned}\right\} \tag{19}$$

在保留到 ε 的同级小量时，(12)式成为

$$W\approx\exp(\mathrm{i}\Gamma'/2) \times R(-\psi_d)\begin{bmatrix} \cos X-\mathrm{i}\Gamma\sin X/2X & \psi_d\sin X/X \\ -\psi_d\sin X/X & \cos X+\mathrm{i}\Gamma\sin X/2X \end{bmatrix}, \tag{20}$$

即节 7.3(4)式，常数相位项 $\exp(\mathrm{i}\Gamma'/2)$在(4)式中已略去.

Chebyshev 等式可表为

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^m = \begin{bmatrix} \dfrac{A\sin m\alpha-\sin[(m-1)\alpha]}{\sin\alpha} & B\sin m\alpha/\sin\alpha \\ C\sin m\alpha/\sin\alpha & \dfrac{D\sin m\alpha-\sin[(m-1)\alpha]}{\sin\alpha} \end{bmatrix}, \tag{21}$$

式中 m 为正整数，矩阵元素 A,B,C,D 满足条件

$$\left.\begin{aligned} C&=-B, \\ AB-BC&=1. \end{aligned}\right\} \tag{22}$$

即矩阵的模为 1，且有

$$\alpha=\arccos[(A+D)/2]. \tag{23}$$

证明如下：把矩阵表为

$$\begin{aligned} Q&=\begin{pmatrix} A & B \\ C & D \end{pmatrix} \\ &=\begin{pmatrix} \cos\alpha-\mathrm{i}\sin\alpha\cos\beta & \sin\alpha\sin\beta \\ -\sin\alpha\sin\beta & \cos\alpha+\mathrm{i}\sin\alpha\cos\beta \end{pmatrix} \\ &=\cos\alpha\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}+\sin\alpha\begin{pmatrix} -\mathrm{i}\cos\beta & \sin\beta \\ -\sin\beta & \mathrm{i}\cos\beta \end{pmatrix} \\ &=\cos\alpha\cdot I+\sin\alpha\cdot J, \end{aligned} \tag{24}$$

式中 I 为恒等矩阵，而

$$J=\begin{pmatrix} -\mathrm{i}\cos\beta & \sin\beta \\ -\sin\beta & \mathrm{i}\cos\beta \end{pmatrix}. \tag{25}$$

且有

$$\left.\begin{aligned} J^2&=-I, \\ J^3&=-J \\ J^4&=I, \end{aligned}\right\} \tag{26}$$

可见矩阵 J 相当于 2×2 矩阵中的“虚数单位”，从而

$$\begin{aligned} &(\cos\alpha\cdot I+\sin\alpha\cdot J)^m \\ &=\sum_{k=0}^{m}\frac{m!}{k!\,(m-k)!}\cos^{m-k}\alpha\cdot\sin^k\alpha\cdot J^k \\ &=\sum_{k=0}^{[m/2]}\frac{(-1)^k m!}{(2k)!\,(m-2k)!}\cos^{m-2k}\alpha\cdot\sin^{2k}\alpha\cdot I \\ &\quad+\sum_{k=0}^{[(m-1)/2]}\frac{(-1)^k m!}{(2k+1)!\,(m-2k-1)!}\cos^{m-2k-1}\alpha\cdot\sin^{2k+1}\alpha\cdot J \\ &=\cos m\alpha\cdot I+\sin m\alpha\cdot J, \end{aligned} \tag{27}$$

在最后一步中用到 $(\cos\alpha+\mathrm{i}\sin\alpha)^m$ 的二项式展开公式，以及 $(\cos\alpha$

$+\mathrm{i}\sin\alpha)^m=\exp(\mathrm{i}m\alpha)$. (27)式可改写为

$$Q^m=\cos m\alpha\cdot I+\sin m\alpha\cdot(Q/\sin\alpha-I\cos\alpha/\sin\alpha)$$

$$=\frac{\sin m\alpha}{\sin\alpha}Q-\frac{\sin(m-1)\alpha}{\sin\alpha}I, \tag{28}$$

即(21)式.

第八章　光学神经网络

8.1 引　　言

本章将研究实现混合神经网络的方案和体系，光学系统主要用于实现矩阵-向量或张量-矩阵的乘法．此外，还将介绍神经网络的基本原理以及全息关联记忆(associative memory)．

随着 80 年代初对于神经网络的兴趣的复苏，Psaltis 和 Farhat 于 1985 年发表了第一篇用光学方法实现神经网络的论文[1]．与光学有关的神经网络称光学神经网络．然而，事实上应当正确地称之为光电混合神经网络，简称混合神经网络 (hybrid neural network)．因为神经网络一般都有一个非线性函数，这是神经网络的重要特征，而用光学方法很难计算这一非线性函数．因此，通常的做法是用电子学方法来实现它．Caulfield 等[2]，Yu[3]，Jutamulia 和 Yu[4]分别发表了关于光学神经网络的综述论文．

神经网络是模拟人脑的结构和功能的信息处理系统．从实际的观点来看，我们并不打算制造一个人工脑，而是设法构造一个信息处理系统去模拟人脑的某些行为．这一系统应当和 von Neumann 型计算机的结构和算法有实质上的区别．光学神经网络可以看作光学信息处理和光计算系统的延伸和扩展，这也是我们将它纳入本书的原因．我们的典型方法是利用已很成熟的神经网络的理论模型，并考虑如何用光学的方案部分地实现它，这一方案在某些方面将比电子学方案更具优越性．

在本章第二节中我们将首先讨论光学神经网络的基本原理，在其次的两节中将分别介绍神经网络的基本概念和实现它的光学方案．在以下的一节中再介绍神经网络的选择学习算法(selected

learning algorithms). 最后将给出一些结论性的评述.

8.2 一般基础

8.2.1 研究神经网络的原动力

神经网络的研究可以追溯到1943年,当时McCulloch和Pitts建立了一个神经活动的理论模型[6]. 在80年代,下面两个因素推动了对神经网络的研究工作:

(1) 尽管计算机可以精确而快速地进行复杂的计算,但在诸如识别一个物体这一类课题上,计算机要末失败,要末比人脑甚至动物的脑慢得太多.

(2) 计算机需要一组精确而复杂的指令来完成某一特定的计算或课题,它活像是一个奴仆,忠实而强健,但绝不聪明. 这样一来,计算机还需要聪明的使用者,他们是高智商的,无需得到详细的命令,就可以指挥奴仆来完成上述任务.

换句话说,一个"聪明的"计算机应当具备这样的功能:一旦某一任务被用户确定后,它能够自动去学习、掌握正确完成该项任务的方法. 上述讨论可归结为"智能型"计算机的两个基本要求:识别能力、学习能力. 这些能力应当通过建立一个模拟人和动物的脑的系统来获得,因为脑同时具备识别能力和学习能力.

8.2.2 神经

人类的脑由大约10^{11}个神经元构成,几乎相当于银河系中恒星的总数. 神经元由三个部分组成:(1)细胞体;(2)树突(dendrites)及(3)轴突(axon). 细胞体包含神经细胞核及生物化学机构,它可以合成细胞核的生存所必需的酶及其他分子. 树突则像灌木丛一般形成许多分叉,围绕着细胞体,神经细胞正是通过它们形成的表面接受输入信号. 轴突则是细胞体的延伸部分,它提供了信号从该细胞体传递到其他细胞的通道.

神经内包含着某种液体，其成分与神经外的液体差别很大. 外部液体所含的钠约为内部液体的十倍，而内部液体所含的钾约为外部的十倍. 由于在细胞的一侧与另一侧钠离子和钾离子(均为阳离子)的浓度不同，轴突中心区域内部与外部约有 -70 mV 的电位差. 当阳离子穿过细胞膜上的通道进入细胞时，内部的电位将逐渐由负变为正，当电位差变为 40 mV 时，另一组通道开启，释放阳离子，使电位差恢复成 -70 mV. 这一电压脉冲称为作用电压或电神经信号. 该电脉冲信号沿着轴突，一直传播到轴突与其他神经元的树突的结合处.

位于轴突的端点和另一神经元之间的结合部是一个突触，在这里轴突膨大而形成终端球状结构，其中包含突触囊. 当电神经信号到达终端球状体时，一些突触囊释放它们的成分——因此又称为化学"发射器"或传递器——它们被另一神经细胞的树突所接收. 信息就这样通过化学传递器一程接一程地在神经元间传输. 接收到信号的神经元，其树突的化学传递器起了开启和关闭细胞膜上的通道的作用，以改变细胞中的离子浓度. 某些突触处于兴奋态时，它们总是要引发神经脉冲；而另一些突触处于抑制态时，它们能够压制或消灭一些信号，使之不可能再激励神经使之发射神经脉冲.

8.2.3 数学模型

上面所描述的神经的行为可以概括如下：

(1) 一个神经元发射一个电脉冲作为它的输出信号. 脉冲的形状永远相同，因而它的输出可以用一个二元状态来表述：1 表示发射，0 表示不动作.

(2) 该电脉冲作为输入信号被另一个神经元所接收. 树突作为化学传递器，接收电脉冲所荷载的信号. 传递器的数目是事先确定的，传递器可以促进或抑制信号的传递. 突触或输入的权重正比于传递器的数目，由整数表示. 突触为正或负，分别表示兴奋

或抑制．如果采取归一化的表示，则突触为实数．

(3) 神经的发射作用，乃是突触被成百上千个神经作用所激活的的综合效果，是一个非线性过程．

综上所述，神经的机理可以用数学公式表为

$$z_j=f\left\{\sum_i w_{ji}x_i+\theta_j\right\}, \tag{1}$$

其中 z_j 是第 j 个神经元的输出，w_{ji} 是第 i 个神经元的输入和第 j 个神经元间的互连权重，x_i 是来自第 i 个神经元的输入，θ_j 是第 j 个神经元的偏置，而 f 则表示某个非线性的传递函数．注意 z_j 和 x_i 都是二元的，取值 1 或 0．在最简单的情况下，f 是阈值函数，表为

$$z_j=1,\quad 当\left\{\sum_i w_{ji}x_i+\theta_j\right\}\geqslant T, \tag{2}$$

以及

$$z_j=0,\quad 当\left\{\sum_i w_{ji}x_i+\theta_j\right\}<T, \tag{3}$$

式中 T 为阈值．当然传递函数也可能是 S 型函数而不是简单的 step 函数．神经元由图 8.1 表示，这一简单的数学模型是由 McCulloch 和 Pitts[6]提出的．

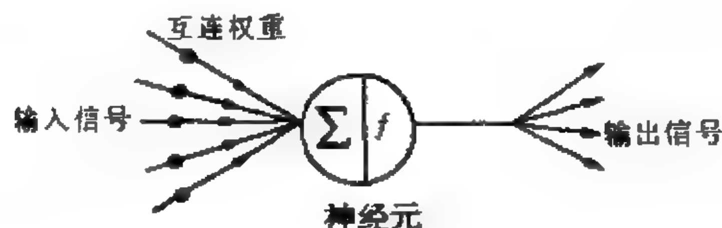


图 8.1　神经元的图示

8.2.4　神经网络

脑是一个由大约 10^{11} 个神经元构成的复杂的网络，每一个神经元通过约 1000 个突触与它周围的神经元发生互连作用，总共有

10^{14}个互连通道. 倘若我们假设神经元只有两个状态:"发出信息"或"不发出信息",则脑的自由度总数为 $2^{10^{11}}$,即 2 的 10^{11} 次幂. 研究表明,每一个思维,每一个感觉都是通过这个网络的活动实现的[8].

由于我们的目的并不在于建立一个人工脑,而在于解决实际的工程问题,因而我们可以把神经网络的问题简化,如图 8.2 所示. 这一"工程神经网络"可如下描述:

(1) 神经元是分层排列的,同一层内的神经元没有相互作用. 任何一个神经元都不与它自身关联.

(2) 每一神经元都按拓扑学的方式与相邻层面上的神经元发生互连,亦即从上一层面上的所有神经元上接收输入信号,并向下一层面上的所有神经元发出信号. 互连权重由实数表示,包括零,正数和负数.

(3) 每一神经元的输出都是二元的.

(4) 信号同时、并行地向一个方向传输,称为前向馈送.

数据则由神经元状态图表达. 可以相信,输出和输入间的任何数学关系,事实上都可以用网络的互连图来表征. 乍一看,似乎将一个数学关系用互连图表示即便不是不可能,也会是相当困难的. 然而,通过一种称为"学习"的特殊方法,数学关系式可以自动地翻译、移植到互联网络中去,这一学习过程的特点在于,通过对于"输入-输出"对的修正,反复调节网络. 应当注意,网络并不能记住输入-输出对,而是在输入和输出之间合成数学关系式,并将合成的关系式存储在互连图案之中. 每当新的输入-输出对参加

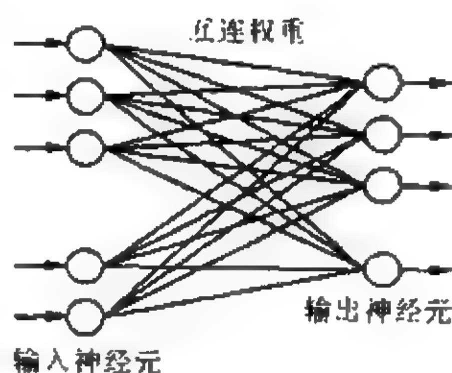


图 8.2　两层神经网络图示

到网络中来，合成关系就获得改进，这一过程称为“训练”．由于该关系式是存储在互连图中，而不是由网络去记住输入-输出对，因而一个从未出现过的新的输入被网络接收时，根据已存储在互连图中的一般关系，该输入信号将引发一个正确的输出信号．

当输入和输出之间的关系模糊，很难描述的时候，这种能力尤其有用．例如，很难写出一组明确的计算机指令，去识别人脸由于感情、视角、阴影的不同而引起的变化．然而，如果给出一系列输入-输出对，其中输入为不同条件下同一个人的脸，而输出则是他的名字，则网络将有能力把不同的面部图像与他的名字关联起来，因为网络在训练过程中已抽取了面部的特征．

我们在原则上已认识到网络满足下面的要求：(1)具有识别能力以及(2)具备学习的能力．这些能力正是我们在神经网络的原动力一节中所讨论的．然而，神经网络并不能在所有的方面都胜过 von Neumann 型计算机．下面的事实是众所周知的：人脑——最优秀的神经网络，在科学计算方面与计算机甚至便宜的计算器相比，却是大为逊色的．

8.3 光学互连

8.3.1 矩阵-向量乘法

上一节方程(1)可改写为

$$z_j = f\{y_j + \theta_j\}, \tag{1}$$

其中

$$y_j = \sum_i w_{ji} x_i. \tag{2}$$

尽管用光学方法实现方程(1)很困难，但用光学系统实现方程(2)却是直截了当的事，这是一个矩阵-向量乘法．而方程(1)可以用电子学方法实现[1]．矩阵-向量乘法的光学实现基于 Goodman 等提出的系统[9]，参见图 8.3(a)，向量 x_i 通过水平放置的 LED 线

阵并行输入系统．透镜 L_1 使每一个 LED 都变成平行光，平行光又被柱面镜 L_2 聚焦成一条线．从每一个 LED 辐射的光束充满矩阵模板 M 的一列，M 上各像素的透过率表征互连权重 w_{ji}．在水平平面（弧矢面）内，M 上每一像素辐射的光束被柱面镜列阵 LA 相应的柱面镜准直．而柱面镜 L_3 对光束不起作用，光束通过透镜 L_4 后会聚到探测器的中心，见图 8.3(c)．在垂直平面（子午面）内，柱面镜列阵 LA 不起作用，柱面镜 L_3 和透镜 L_4 构成 1：1 成像系统，将 M 上的每一行聚焦到探测器的相应单元上，这些单元对 M 的一行像素进行积分，给出矩阵-向量乘法的结果 y_j，见图 8.3(b).

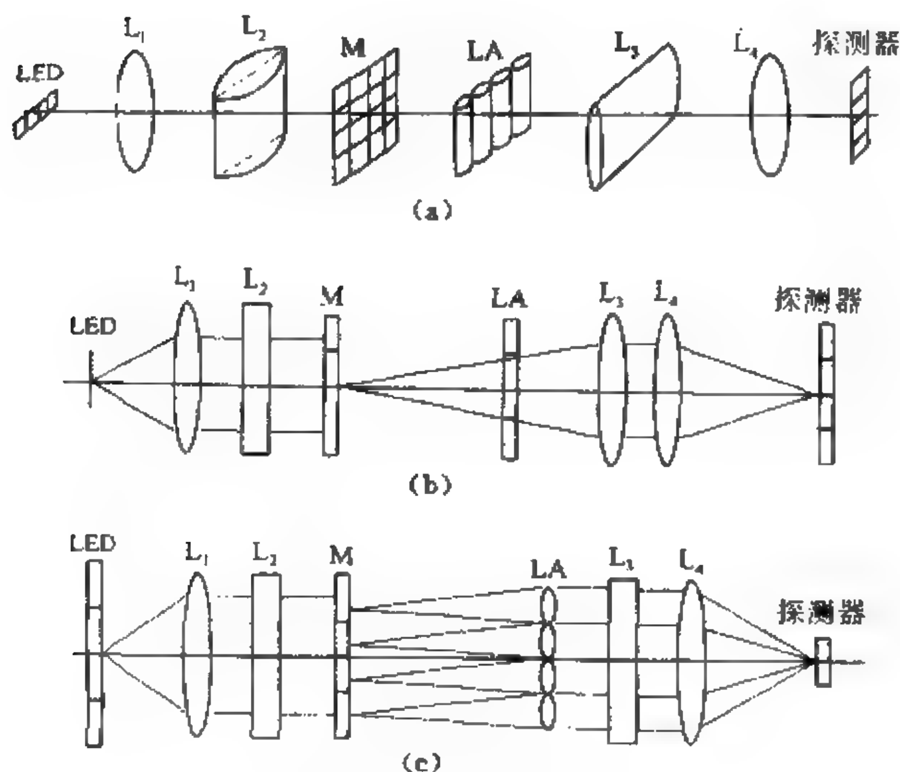


图 8.3　非相干乘法器

非相干光处理器的主要问题在于，它不能同时处理正数和负数. 因此，上述矩阵-向量乘法器仅用于单极向量和矩阵. 在神经网络中，向量都是二元的，其元素由 0 和 1 表示. 由于矩阵通常是单极的，我们需要在矩阵-向量乘法器中有两个通道，分别对应于正和负的数值. 最后正和负的结果可以用电子学的方法加以联合，正如光学神经网络的最早的一篇论文所指出的[1]，正、负通道可以用一对正交的偏振态来编码[10,11]，也可以用不同时间的帧图像来编码[12]. 另一做法是，在每一权重中增加一个偏置，并相应调整阈值[13,14]. 也有人证明，当略去负的权重时神经网络的性能反而提高[15].

然而，如果我们应用相干光处理系统（下面还会详细讨论），原则上只需用简单的处理器，因为被处理的参数是复振幅而不是强度，振幅可正可负，甚至可以是复数. 人们对复互连网络的行为很有兴趣，尽管它尚未普遍使用[16].

在另一方面，上述矩阵-向量乘法器的一个优点正在于它以非相干模式工作，调制参量和探测参量相同，都是光强. 而在相干光处理系统中，调制参量是振幅，探测参量是光强，这一差异在一些用相干光系统实现的神经网络中常常被忘记. 探测到的光强度是矩阵-向量相乘的结果 y_j 的平方，而 y_j 的值可以通过电子学的方法加以恢复，例如用计算机来恢复. 采用相干光后，图 8.3 所示的系统可大为简化，这是因为一个扩展的非相干面光源不可能用光学方法聚焦成一个小点，这从测不准关系就可以预见. 然而被一个点光源（例如相干光）照亮的面积可以重新聚焦成一个小点.

相干光矩阵-向量乘法器如图 8.4 所示. 相干光首先由透镜 L_1 准直，准直光通过矩阵掩模板 V，它由若干长条形区域构成，每一列代表一个 x_i，然后通过矩阵模板 M，它包含权重函数 w_{ji}. 通过同一行的光线由柱面镜 L_2 聚焦到一个垂直放置的一维探测器列阵的一个单元上，每一单元探测到的信号表征矩阵-向量积的一个分量，这一系统是 Kranzdorf 等[10]提出来的.

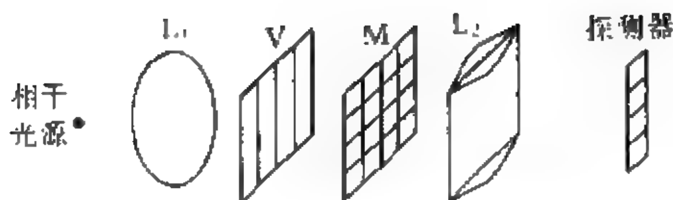


图 8.4　相干光矩阵-向量乘法器

矩阵-向量乘法器甚至可以不用透镜而设计得非常紧凑. 线状非相干光源列阵和线状探测器与一个矩阵掩模板叠合使用就实现了这一乘法运算,见图 8.5,线状光源均匀地照射到模板的一列上,而线状探测器则对一行的光强求和. 由于这一结构非常简单紧凑,在许多混合光学神经网络单元中得到运用[17~20].

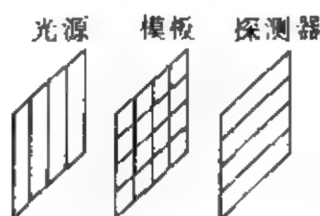


图 8.5　简化的矩阵-向量乘法器

8.3.2　透镜列阵

尽管作者还不清楚人的大脑是否把二维信息保持在相同的二维神经图案中,但图 8.2 所示的一维工程神经网络却容易扩展,以处理二维信息. 相应地,(1)、(2)式变成

$$x_{jl}=f\{y_{jl}+\theta_{jl}\},\tag{3}$$

其中

$$y_{jl}=\sum_i\sum_k w_{jlik}x_{ik},\tag{4}$$

这里的输入输出都是二维矩阵,互连则是四维张量. 与(2)式相似,(4)式也可以由光学系统实现.

为了更好地理解透镜列阵的作用,我们由(2)式所表示的矩阵-向量乘法入手. 假设要处理的为 3×1 向量及 3×3 矩阵,则每一个向量元素表示如下:

$$y_1=w_{11}x_1+w_{12}x_2+w_{13}x_3,\tag{5}$$

$$y_2 = w_{21}x_1 + w_{22}x_2 + w_{23}x_3, \tag{6}$$

$$y_3 = w_{31}x_1 + w_{32}x_2 + w_{33}x_3, \tag{7}$$

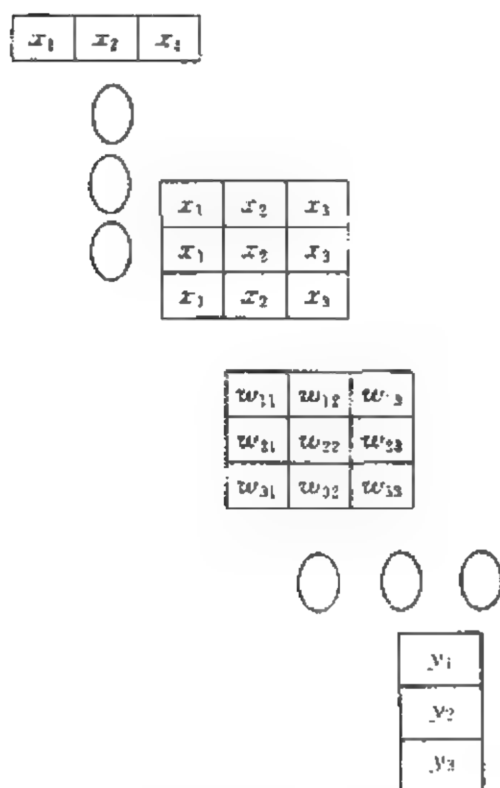


图 8.6 用两个透镜阵实现矩阵-向量乘法运算

方程(5)～(7)可以用两个互相正交的透镜列阵实现，见图 8.6，第一个列阵把向量(x_1, x_2, x_3)成像到矩阵模板w_{ji}上去，第二个列阵则使同一行内的元素移位并叠加，形成结果向量(y_1, y_2, y_3)，为了简单起见，我们忽略了透镜成像过程中的 180°旋转，本章中我们将始终忽略像的旋转. 与图 8.3、图 8.4 对比，可以看出我们现在并不需要柱面镜.

图 8.6 所示的光学系统可以加以扩展，以处理各种二维问题. 为处理上的简化，我们设(4)式中要处理的是一个 2×2 矩阵. 结果的矩阵中的元素由下式给出：

$$y_{11} = w_{1111}x_{11} + w_{1112}x_{12} + w_{1211}x_{21} + w_{1212}x_{22}, \tag{8}$$

$$y_{12} = w_{1121}x_{11} + w_{1122}x_{12} + w_{1221}x_{21} + w_{1222}x_{22}, \tag{9}$$

$$y_{21} = w_{2111}x_{11} + w_{2112}x_{12} + w_{2211}x_{21} + w_{2212}x_{22}, \tag{10}$$

$$y_{22} = w_{2121}x_{11} + w_{2122}x_{12} + w_{2221}x_{21} + w_{2222}x_{22}. \tag{11}$$

参见图 8.7，一个透镜把输入矩阵模板 x 成像在互连张量模板 w 上，w 已进行分块处理，并构成二维模板的形式. 后面的一个透镜列阵把 x 的像和 w 进行叠加，形成输出矩阵 y. 对于 $N\times N$ 的输入矩阵，互连张量的元素为$(N\times N)^2$ 个，而列阵中的透镜数也等于 $N\times N$.

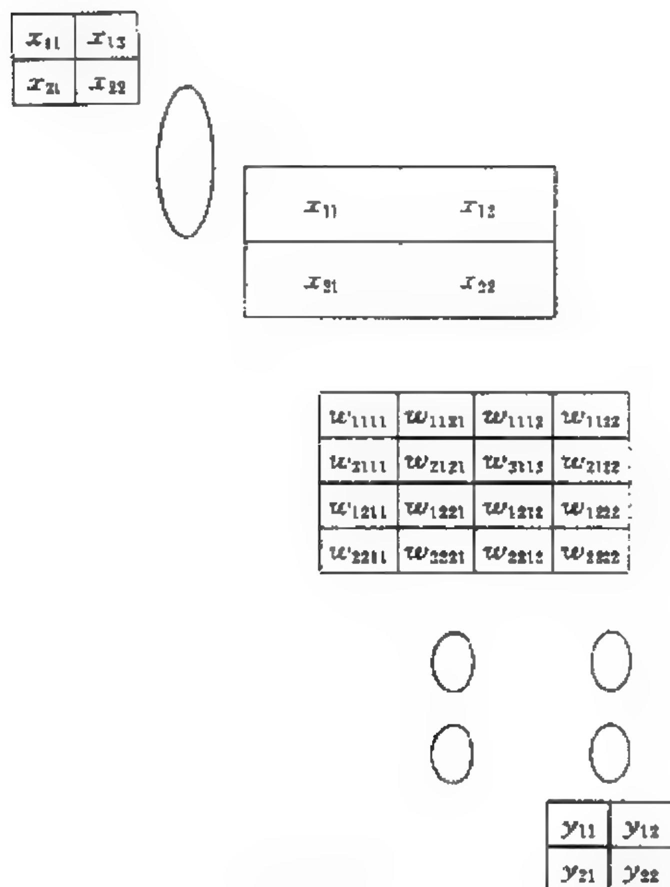


图 8.7 用成像透镜和透镜列阵实现张量-矩阵乘法

Farhat 和 Psaltis 则提出另一种构想[17]. 他们用透镜列阵复制输入矩阵并成像在互连张量 w 上，如图 8.8 所示，一个面探测器对 w 的每一部分分别积分. 注意图 8.7 和 8.8 中 w 像素的安排不相同. 图 8.7 走的是图 8.6 的路线，后者又是图 8.3 方案的简化形式. 然而图 8.8 中的张量 w 的分块和重组方案比较好，且每一个面探测器可以用一个积分透镜和一个点探测器代替，探测器位于透镜焦点上(见图 8.9). 除非用的是相干光，否则严格说来，透镜并不能对张量的各部分进行积分，它所探测的则是这一部分的光强.

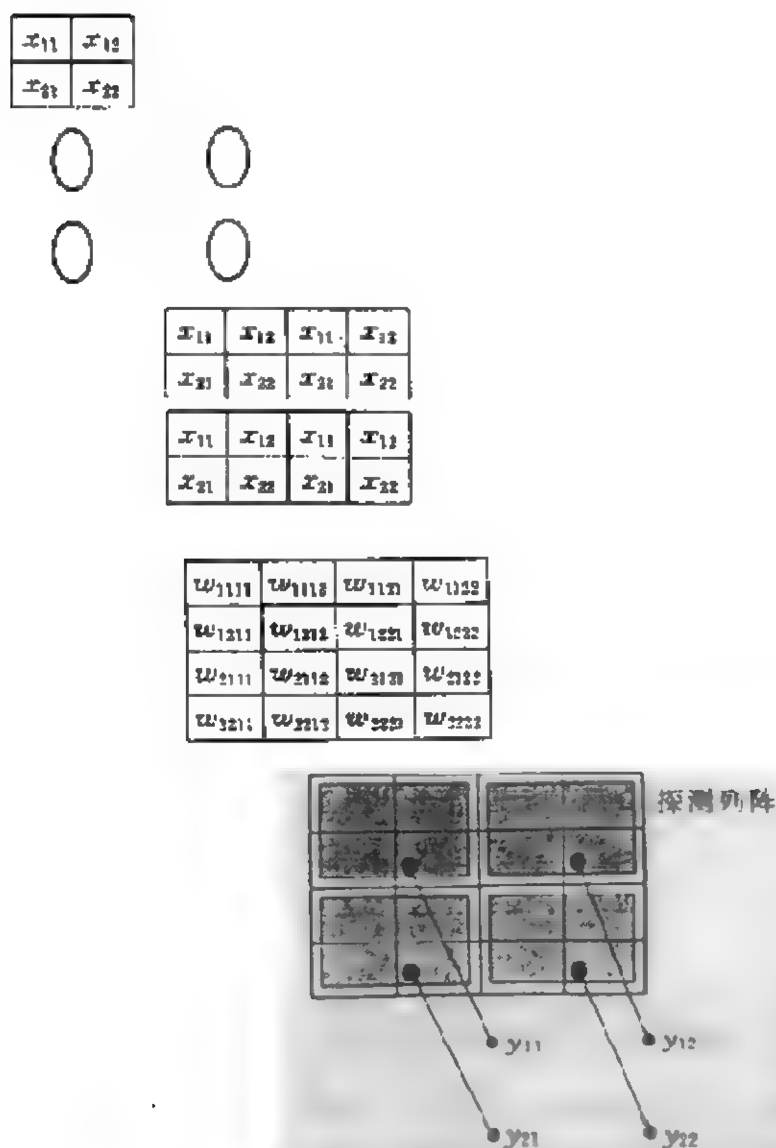


图 8.8 用透镜列阵和面探测器实现张量-矩阵乘法

Wu 等则提出另一种机构[21]，在其中分块张量 w 放在输入矩阵 x 前，如图 8.10 所示. 透镜 L_{11} 把 $w_{1111}, w_{1112}, w_{1211}, w_{1212}$ 分别成像到 $x_{11}, x_{12}, x_{21}, x_{22}$ 上，如此等等. 一个单透镜再把透镜列阵中的各单元透镜 L_{11}, L_{12}, L_{21} 和 L_{22} 分别成像到面探测器 $y_{11}, y_{12}, y_{21}, y_{22}$ 上去，它们的输出给出矩阵-张量相乘的结果. 其实用小孔列阵代替透镜列阵也能给出同样的结果，请参见 Lin 等的论文[22].

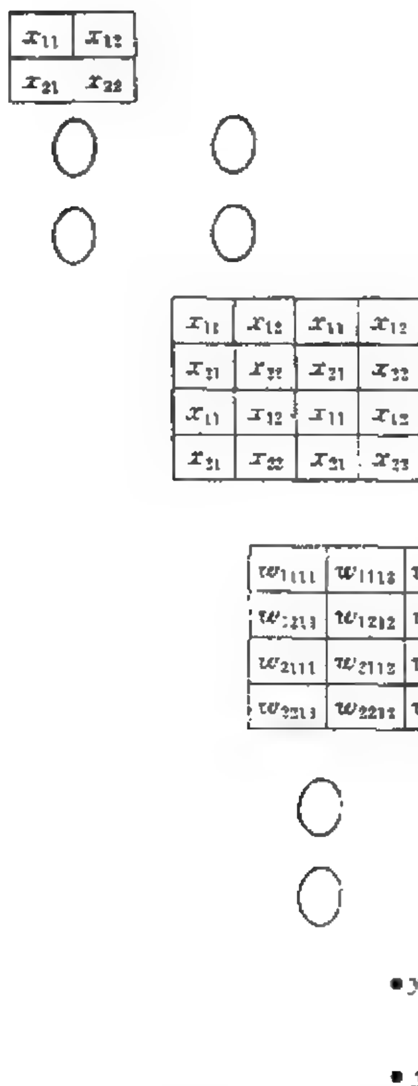


图 8.9　用两个透镜列阵实现张量-矩阵乘法

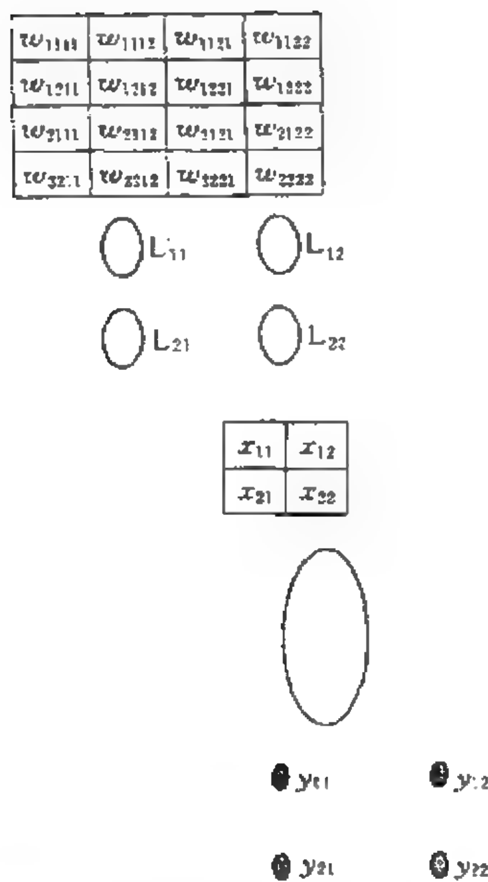


图 8.10　应用透镜列阵和成像透镜实现张量-矩阵乘法

通常张量 w 的动态范围很大，不能用 SLM 来表征 w. 一个很有意思的设计是由 Wu 等提出来的[21]，他们用动态范围很大的计算机监视器(montior)来显示互连张量，而 SLM 则用作输入矩阵的显示，因为输入信号是二元的. 此外，透镜列阵还可以用特殊设计的全息光学元件(HOE)或衍射光学元件(DOE，即二元光学元件 BOE)来代替[23].

8.3.3 光学相关器

分块张量可表为矩阵 w_{11}，w_{12}，w_{21}和 w_{22}. 方程(8)～(11)表明，y_{11}是输入矩阵 x 和矩阵 w_{11}的内积，y_{12}则是同样的输入矩阵 x 和 w_{12}的内积，等等. 由于内积表达式相当于在原点算出的相关函数，所以公式(8)～(11)也可以用相关来完成. 图 8.10 明确表示 y_{11}是输入矩阵 x 和矩阵 w_{11}在零点的相关，y_{12}是输入矩阵 x 和矩阵 w_{12}在原点的相关，等等.

最近，Jutamulia[26]提出用自适应联合变换相关器来实现(8)～(11)，也就是(4)式，参见图 8.11. 联合变换相关(JTC)原是 Weaver 和 Goodman[27]、Rau[28]独立提出来的. 两个在空间分离

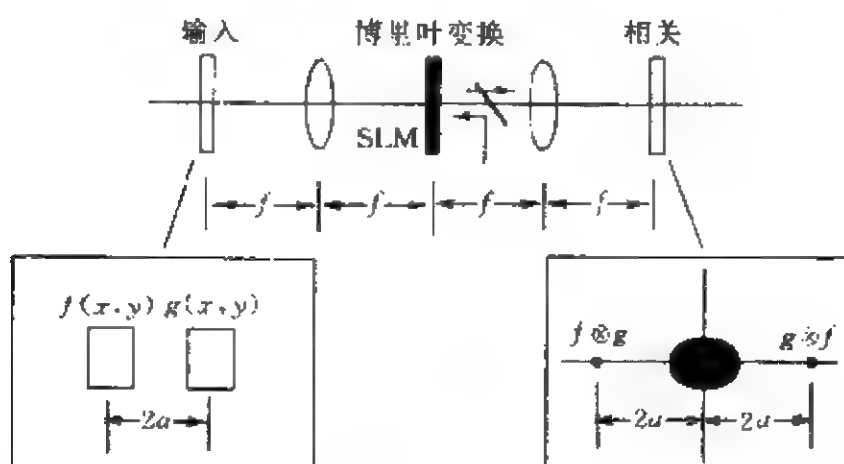


图 8.11 Joint 变换相关器示意图

的函数 $f(x+a,y)$和 $g(x-a,y)$放在一个相干光傅里叶变换器的输入平面上，用相干的平面光波照明，第一个透镜给出 f 和 g 的傅里叶变换，表为 F 和 G，它们在频谱面的中心相互叠加，形成复振幅和 $F+G$，叠加后的光强则为$|F+G|^2$. 一个空间光调制器 SLM 将光强分布转换到读出光的振幅函数中，经过第二个透镜，$|F+G|^2$发生傅里叶变换得到

$$\mathscr{F}\{|F+G|^2\}=\mathscr{F}\{|F|^2+F^*G+FG^*+|G|^2\}$$
$$=\mathscr{F}\{|F|^2\}+\mathscr{F}\{|G|^2\}+\mathscr{F}\{F^*G\}+\mathscr{F}\{FG^*\},\tag{12}$$

其中前二项分布在输出平面的中心附近，而

$$\mathscr{F}\{F^*G\}=f\otimes g,\tag{13}$$

式中⊗表示相关. 这一结果的推导可参见节 1.1，或本章文献[29]. (12)式的后二项则是 $f(x+a,y)$ 和 $g(x-a,y)$ 的相关及其共轭像，也就是 $f(x,y)$ 和 $g(x,y)$ 的相关从原点沿 x 轴位移 $-2a$，以及它的共轭像沿 x 轴位移 $2a$ 的结果.

为了实现(8)～(11)式，输入矩阵模板 x 及分块矩阵模板 w 放置在联合傅里叶变换相关器的输入平面上，可参见图 8.12，

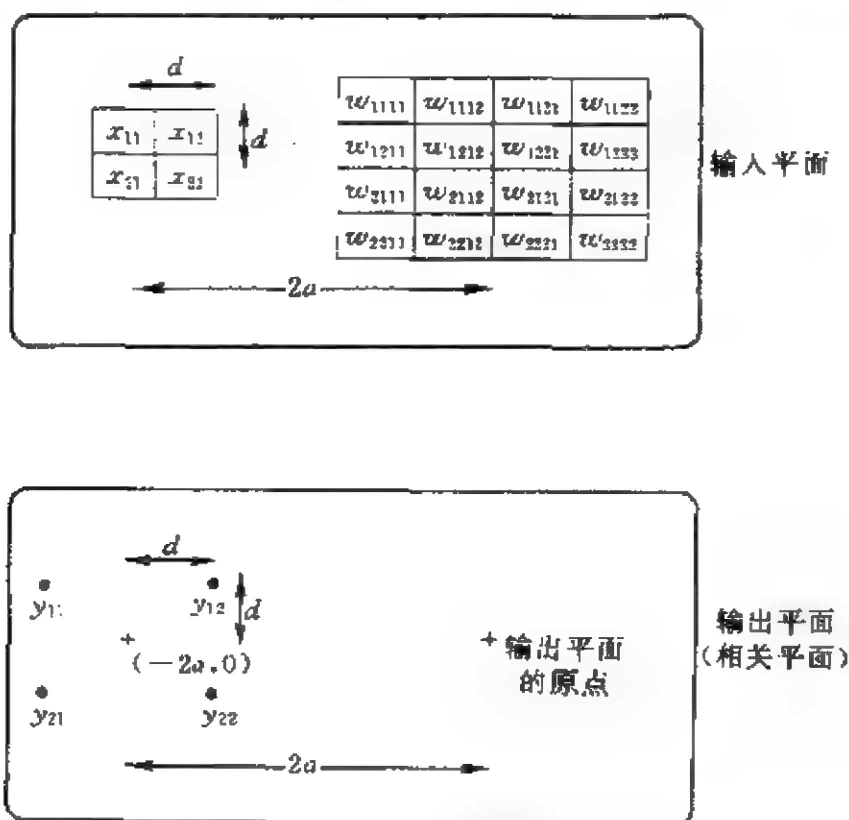


图 8.12 使用 JTC 来实现张量-矩阵乘法

x 与 w 的中心相距 $2a$，每一像素的尺寸为 $d\times d$. x 和 w_{11}，w_{12}，w_{21}，w_{22}的相关分别位于输出平面的相应位置，形成 y_{11}，y_{12}，y_{21}，

y_{22}. 应用JTC的优点在于:(1)透镜列阵加工困难较大,特别是当矩阵 x 较大时,加工更困难. JTC的应用代替了透镜列阵. (2) JTC具备表达单极性互连权重的能力. 然而,正由于JTC是相关光学处理器,测到的强度是光学计算值的平方.

作为JTC的孪生兄弟,Vander Lugt相关器[30]也可用于实现(8)~(11)式,请参见Chao和Stoner[31]的文章. 与JTC不同,Vander Lugt相关器中滤波器是很难调试的,当然也可以用光折变晶体来生成实时的匹配滤波器[32,33]. Chao和Stoner[31]仍用透镜列阵来实现多通道处理,而Psaltis等[32],以及Hong等[33]则用多角度体积全息来实现多通道处理.

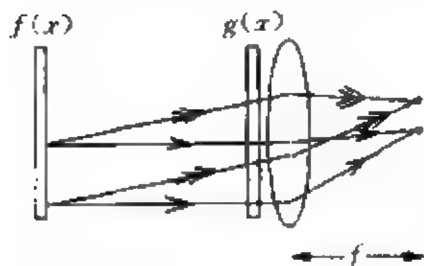
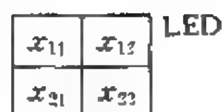


图8.13 投影相关示意

Lin等[22]早在1989年就指出了相关处理的途径,他们使用了非相干投影方法,参见图8.13. $f(x)$上每一像元都向各个方向辐射一个非相干光波,该光波透过函数 $g(x)$掩模板上所有的像元,同一方向的光束将会聚在透镜后焦面上的同一点. 焦面的光强分布则提供了 $f(x)$和 $g(x)$的相关. 为了模拟(8)~(11),用四个LED来表示 x_{11},x_{12},x_{21}和 x_{22},分块矩阵模板则安排在LED后面,如图8.14. 由于通过 $x_{11}\sim w_{1111}$,$x_{12}\sim w_{1112}$,$x_{21}\sim w_{1211}$和 $x_{22}\sim w_{1212}$的光线具有相同的方向,它们经过透镜聚焦在同一点 y_{11},通过 $x_{11}\sim w_{1121}$,$x_{12}\sim w_{1122}$,$x_{21}\sim w_{1221}$和 $x_{22}\sim w_{1222}$的光线也具有相同的方向,经透镜后聚焦在 y_{12},等等. 各种光学相关器[36]都可用于实现神经网络.

8.3.4 全息术方法

全息图记录物光和参考光形成的干涉图. 参考光通常从一个点源辐射. 当用参考光再次照射全息图时,就可以用光学方法重构原来的物体. 在全息技术中,分块矩阵模板可以用一组全息图



w_{1111}	w_{1112}	w_{1121}	w_{1122}
w_{1211}	w_{1212}	w_{1221}	w_{1222}
w_{2111}	w_{2112}	w_{2121}	w_{2122}
w_{2211}	w_{2212}	w_{2221}	w_{2222}





图 8.14　用投影相关法实现张量-矩阵乘法

来记录．当用输入矩阵 x 重构时，方程(8)～(11)就可以用全息方法有效地实现[13,37]，参见图 8.15.

x_{11}	x_{12}
x_{21}	x_{22}

H_{11}	H_{12}
H_{21}	H_{22}

y_{11}	y_{12}
y_{21}	y_{22}

图 8.15　用全息术实现张量-矩阵乘法运算

首先，我们用 x_{11} 作为参考光来制作分块矩阵 w_{11}' 的全息图 H_{11}. 当我们用位于 x_{11} 的点源照射全息图 H_{11} 时，在输出矩阵 y 处产生分块矩阵. 即在 y_{11} 处生成 w_{1111}；在 y_{12} 处生成 w_{1121}；在 y_{21} 处生成 w_{2111}；在 y_{22} 处生成 w_{2121}，请与图 8.7 相对比. 重构的 w_{11}' 的强度正比于点源 x_{11} 的光强. 用同样的方法，由 w_{12}' 和 x_{12} 制作 H_{12}；用 w_{21}' 和 x_{21} 制作 H_{21}；用 w_{22}' 和 x_{22} 制作 H_{22}. 与图 8.7 的系统类似，图 8.15 中 y_{11}～y_{22} 处的强度即表征(8)～(11)式.

Ivanov 等把一组一维全息图装在一个可旋转的光学转盘系统上以实现矩阵-向量乘法[38]. 在时刻 t_1，输入 x_1 通过一维全息图在 y_1，y_2 和 y_3 处分别产生 $w_{11}x_1$，$w_{21}x_1$ 及 $w_{31}x_1$. 在时刻 t_2，由于转盘转到另一位置，输入 x_2 通过全息图在 y_1，y_2 和 y_3 分别产生 $w_{12}x_2$，$w_{22}x_2$，和 $w_{32}x_2$，如此等等. 在时刻 t_1 到 t_3，位于 y_1～y_3 的每一探测器分别积分以得到(5)～(7).

Psaltis 和他的合作者[39,40]用单个体全息图代替分离的全息图列阵来计算(8)～(11). 他们设想，从 x_1 辐射的光波经过体全息图中与互连权重 w_{1111} 对应的光栅的衍射，应当射向 y_{11}，等等. 在(8)～(11)式给出的例子中，有 8 组光栅共存于体全息介质中. 必须指出，光栅 w_{1111} 不会干扰除 x_{11} 以外其他光源辐射的光波[41]. 如果全息光栅 w_{1111} 是由 x_{11} 辐射的光波和另一个会聚于 y_{11} 的光波形成的，如图 8.16，那么只有 x_{11} 辐射的光波才满足 Bragg 条件，

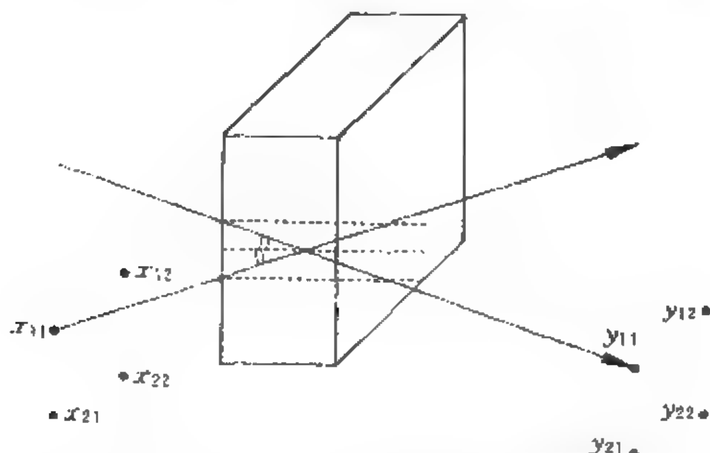


图 8.16　用单个体全息实现张量-矩阵乘法运算

它被 w_{1111} 衍射，衍射波会聚于 y_{11}. 来自其他光源的光波均不满足 Bragg 条件，因而不受 w_{1111} 的影响.

8.3.5 光纤网络

公式(5)～(7)及(8)～(11)所要求的互连可以直接由光纤系统实现. Arrathoon[42]，Goutzoulis[43]，Ito 和 Kitayama[44] 等先后提出各种光纤网络. 但是，除了信号是在光纤中以光波传递这一点以外，该方案与电路方案没有什么不同. 而 Arrathoon[42] 指出光纤网络的扇入和扇出能力比电子可编程逻辑列阵(electronic programmable logic arrays, PLAs)高出 100 至 1000 倍.

8.3.6 光学互连的优点

我们已表明如何用光学系统来实现神经网络中的矩阵-向量乘法运算和张量-矩阵乘法运算，我们尚未讨论过光学互连的优点. 一般来讲，不用超大规模集成电路(VLSI)系统而采用光学互连方案的原因在于：

(1) 电信号在金属互连网络中的传输取决于单位长度的电容. 当互连通道数增大时，传输速度会降下来，而光学互连的速度与互连通道数无关；

(2) 光学信号在空间可以独立传播，彼此间不相互干扰；

(3) 光学信号可在三维自由空间中传输；

(4) 互连可以通过 SLM(空间光调制器)适当地改变；

(5) 光学信号非常容易转变成电信号.

神经网络所要求的矩阵-向量或张量-矩阵乘法是非常特殊的. 此外，在各种用光学系统实现互连的方案中，内积方法较容易实现，因此比较常用，以实现各种矩阵运算[46,47].

在神经网络中采用光学方法的主要原因，在于超大规模集成电路不可能实现高达 10^3～10^6 的互连或扇入-扇出，因为电路中每一对输入-输出都要求单一的分立通道. 由此我们也可以想像，将

来到了每个神经元的互连高于1000时,光学神经网络才真正有了用武之地.

8.4 神经网络的学习

神经网络的特点就在于它具有识别和学习的能力. 神经网络的识别和计算,在一维情况下可通过求解节8.2(1)~(3)式来实现,在二维情况下则可通过求解节8.3(3)式及有关的方程来实现. 前面已讲过,这些操作可以借助于各种混合光学处理系统来进行. 然而,进行这些操作的前提条件,是输入矩阵和互连张量是已知的. 在下面的讨论中,我们将看到如何通过学习的方法来建立这一张量. 建立互连张量的过程称为学习(learning). 这一学习的过程则可运用在线或不在线的计算机来进行,也可通过实施矩阵-向量及张量-矩阵乘法的混合光学处理系统来进行. 为简单起见,我们只做一维情况下的理论分析,这一分析可直接推广到二维.

8.4.1 感知(perception)

由Rosenblatt在1958年提出的神经模型[48]把早期的学习手续归结为感知的或感觉的学习(perception learning algorithm), McCulloch和pitts[6]则运用图8.1来表示这一模型. 由于输出是二元的,神经元只能作出YES或NO的决定. 换言之,神经元可以将输入分成A,B两类,分别用+1或-1来表示,或用1和0来表示. 对于+1/-1表象或1/0表象还需专门加以比较[24].

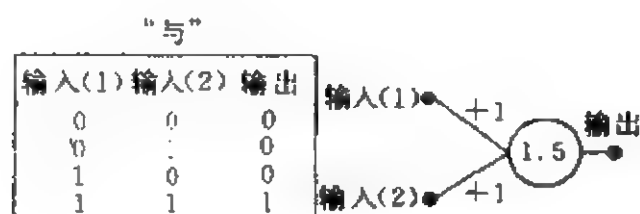
在许多情形下,包含一个神经元的感知模型可表为

$$z=1, \quad 若 \quad \sum_i w_i x_i \geqslant T, \tag{1}$$

或

$$z=0, \quad 若 \quad \sum_i w_i x_i < T, \tag{2}$$

在这里我们把 $T-\theta$ 综合成 T，即神经元的阈值. 为了说明问题，取 $w_1=w_2=1, T=1.5$，则感知功能可表现为对两个输入的综合结果的识别，其结果相当于逻辑“与”(AND)操作，见图 8.17. 输入端常不止两个，但输出只有一个. 在图 8.17 给出的例子中，权的选择是通过“试探-误差”(trial-and-error)即逐次逼近过程进行的. 事实上，感知学习算法也是基于逐次逼近的途径. 首先，权 w_i 和阈值 T 都取为小的随机数. 用一组输入(在图 8.17 中，一组输入(1)和输入(2))及已有的 w_i, T 来计算输出 z. 实际的输出 z 与所



“与”

输入(1)	输入(2)	输出
0	0	0
0	1	0
1	0	0
1	1	1

图 8.17　实现“与”运算的神经网络

要求的输出(目标值)t 进行对比，确定权的改变量如下：

$$\Delta w_i = \eta(t-z)x_i, \tag{3}$$

其中 η 是控制学习过程的增益因子，定义为

$$0<\eta\leqslant 1. \tag{4}$$

根据学习的原则，经过修正的权为

$$w_i = w_{i,\text{old}} + \Delta w_i, \tag{5}$$

这一过程一直持续到

$$(t-z)=0 \tag{6}$$

为止.

另一对输入-输出的学习又开始了. 到全部输入-输出对都试验完毕，学习过程便告结束. 必须注意学习方法中并不包含对阈值 T 和增益因子 η 的确定，这两个参数还是需要猜测或逐次逼近的过程来确定.

理论上可以证明[49]，如果输入的组合是线性可分离的话，权因子 w_i 的解总是存在的．换言之，经过有限的步骤，迭代过程总是以给出正确的互连权重而告终结．然而，如果输入不满足线性可分离条件，则感知学习过程将以迭代手续不收敛而告失败．逻辑“或”(OR)操作是经典的例子，它要求 00→0，01→1，10→1 及 11→0，而且不满足线性可分离的条件．我们不可能用图 8.17 所示的神经元来实现“或”门．如果我们用上述的感知学习手续，迭代过程将不会结束．倘若我们对神经网络的一个新的输入-输出对进行训练的话，原来的输入-输出对则将被放弃．

我们再重复一遍，识别和学习是神经网络的两大特点．我们用电寻址 SLM 作为光学模板表达矩阵和分块张量(如上节所述)，则感知的识别步骤就可以在以前所描述过的混合光学处理系统的框架中顺利实现．$\sum_i w_i x_i$ 的计算由光学进行，它的强度由探测器检测并转换成电信号．而取阈值则运用电子学方法或计算机来实现．由(3)～(6)式表示的学习过程也由电子学方法或计算机执行．计算机算出的新的权 w_i 通过 SLM 代替旧的 w_i，然后由光学系统继续计算下一个 $\sum_i w_i x_i$，等等．

如果用光折变晶体而不是用 SLM 来表示分块张量，学习过程或它的一部分是可以用光学方法实现的，详见 Hong 等[33]的文章，下面我们简单描述一下这一过程．

图 8.18 给出光学感知的示意图．首先，利用与光轴夹角为 α 的参考光，在光折变晶体中制作一个 Vander Lugt 型匹配滤波器，SLM 中给出初始的互连权重 w_{ik}，当匹配滤波器在晶体中建立后，SLM 显示输入 x_{ik}，而 x_{ik} 和 w_{ik} 的相关峰则在离轴 α 角的方向上出现，表征 $\sum\sum_{ik} w_{ik} x_{ik}$．用探测器探测出相关斑的强度，并转换成电信号，再用电子学方法将这一数值与阈值 T 对比，如该数值比阈值大，输出 $z=1$，否则 $z=0$．$t-z$ 为偏差信号，t 是期望的输出．由于 t 和 z 仅取 1 或 0，$t-z$ 可能取 +1，−1 或 0.

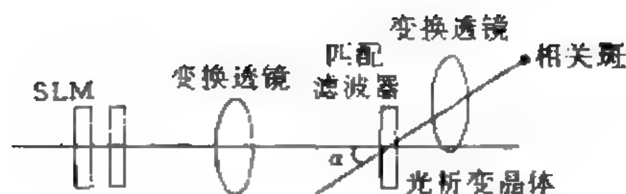


图 8.16 光学感知示意图

如果偏差信号 $t-z$ 为 0，权重已经合适，不需再进行下去．若 $t-z$ 为 1 或 -1，权重必须修正．设 $\eta=1$，由(3)式，有

$$\Delta w_{ik}=x_{ik}, \qquad 若\ t-z=+1; \tag{7}$$

以及

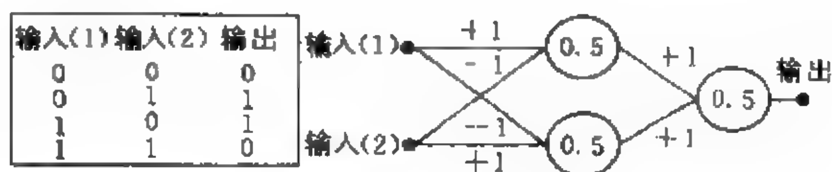
$$\Delta w_{ik}=-x_{ik}, \quad 若\ t-z=-1; \tag{8}$$

若将新的滤波器叠加在旧的滤波器上，互连权重可以有效地修正．若偏差信号为$+1$，将 x_{ik}通过 SLM 输入系统以制作新的附加匹配滤波器；若偏差信号为-1，则在输入 x_{ik}时，在前面附加一个 π 的相位片．当神经元不限于二元值、变换函数不限于取阈值时，(3)式称为最小均方根算法，即 delta 法则，或 Widrow-Hoff 算法[49,50]．McAulay 等[51]提出另一类光学感知方案．Itoh 等[52]提出了应用光致发光材料来实现 delta 法则的学习．Fisher 等[53]及 Kranzdorf 等[10]则用光学或电寻址的 SLM 来进行 Widrow-Hoff 型学习．

8.4.2 反向传播

前面我们已讨论过，感知仅能实现线性可分离的输入的学习，例如“与”．当遇到“或”这一类问题时感知学习便告失败．然而，如果我们把若干个感知组合成分层神经系统，这一多层神经系统就能解决诸如“或”的问题．事实上，多层神经网络可以解决所有的问题．例如，一个两层的神经网络就能根据“或”的法则恰当地

将输入变换为所要求的输出，如图 8.19 所示. 显然感知型学习机构无法用于寻找正确的互连权重，因为对于中间层（又称隐蔽层），我们并不知道所要求的参数的值应当是多少.



输入(1)	输入(2)	输出
0	0	0
0	1	1
1	0	1
1	1	0

图 8.19　实现"或"的神经网络

在反向传播学习机构中，基于神经网络输出端的偏差信号，可以计算出输出层的前一层的偏差信号. 换句话说，为了确定隐蔽层的偏差信号，输出层中的偏差信号被反向传播到输入层去. 利用每一层的偏差信号，权重可以根据类似于(3)式的广义 delta 法则予以修正.

能够进行反向传播学习的最简单的神经网络可如下构造：第 j 层的神经元的输出为

$$z_j=f(y_j), \tag{9}$$

其中 $f(\cdot)$ 并不是阈值函数，而是如下的一个非线性变换函数：

$$f(y_j)=\{1+\exp[-(y_j+\theta_j)]^{-1}, \tag{10}$$

θ_j 为偏置，且有

$$y_j=\sum_i w_{ji}x_i, \tag{11}$$

其中 x_i 是第 i 层的神经元的输出，该层在第 j 层前面，w_{ji} 为互连权重.

与(3)式类似，修改权重的法则为

$$\Delta w_{ji}=\eta\delta_j x_i, \tag{12}$$

输出层的偏差信号为

$$\delta_j=(t_j-z_j)f'(y_j), \tag{13}$$

这里 t_j 是要求的输出，z_j 则是实际输出，且有

$$f'(y_j)=\frac{\partial f(y_j)}{\partial y_j}.\tag{14}$$

代入(10)和(9)式分别得到

$$f'(y_j)=f(y_j)[1-f(y_j)].\tag{15}$$

及

$$f'(y_j)=z_j(1-z_j).\tag{16}$$

当该层不是输出层时,反向传播信号按下式计算:

$$\delta_i=f'(y_j)\sum_j\delta_j w_{ji},\tag{17}$$

其中 i 层先于 j 层,这一手续使偏差向回传播一层,并可对每一层重复上面的手续.

运用上一节的方案,可以在一维和二维情况下用光学方法实现(11)式. 例如,Jutamulia 提出用 JTC 的方案实现(11)式及(17)式中矩阵-向量乘法部分,而用计算机完成其余的非线性运算[55]. Peterson 等则提出用光折变晶体进行自适应互连、用电子学方法进行非线性运算的方案来实现反向传播学习[56]. 用光学方法完成学习过程中的非线性部分仍然是很困难的,尽管 Caltech(美国加州理工学院)的一些研究小组曾提出过一些设想[40,57,58].

图 8.20 给出一个用于混合计算-光学型神经网络的自适应

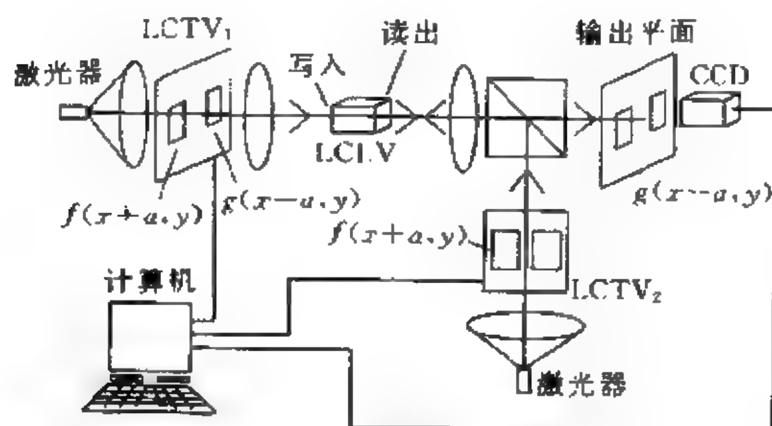


图 8.20　用于混合计算-光学型神经网络的自适应联合变换相关器(JTC)

JTC 的示意图，液晶电视($LCTV_1$)显示一个位于($-a$,0)的透射函数 $f(x,y)$以及位于(a,0)的另一个透射函数 $g(x,y)$. $LCTV_1$ 被一个准直的激光束照明，而 $f(x+a,y)$和 $g(x-a,y)$通过透镜进行联合傅里叶变换，变换的谱函数被液晶光阀(LCLV)转换成功率谱. 如果这一功率谱被一个平面波读出的话(即移去图8.20中的 $LCTV_2$)，在输出平面上将出现 $f(x,y)$和 $g(x,y)$的相关，位于输出平面($-2a$,0)和($2a$,0)附近. 然而，如果 $LCTV_2$ 显示 $f(x+a,y)$，则输出平面上将出现 $g(x-a,y)$.

一个神经网络可以看成是一个普遍变换器，它通过同一互连网络把一组输入同时变换成特定的输出. 一个有监视的神经网络(supervised neural network)例子就是反向传播型学习，在学习期间，互连权重开始时是被随意指定的，然后网络就对某一对输入-输出进行试验. 试验的输入馈入网络，经过加权求和及非线性变换后产生了输出. 与期望的输出相比较后，计算出偏差，其中包含导数计算. 这一学习过程显然是递降的，互连权重根据算出的偏差进行修正，这一过程一直进行到网络产生所要求的结果才结束. 网络再对其他输入-输出对实施同样的试验.

图 8.20 所示的系统适用于前向、后向网络，而光学信号并不需要反向. 在前向传播时，$f(x+a,y)$和 $g(x-a,y)$由计算机生成并显示在 $LCTV_1$ 上. $f(x-a,y)$通过 $LCTV_2$ 反向输入系统去读取 LCLV，在输出平面上就得到 $g(x-a,y)$. 这样一来，$f(x+a,y)$被变换成 $g(x-a,y)$并由 CCD 探测. 对测出信号的非线性变换则由计算机完成. 第二个互连的计算用同样的自适应 JTC 进行.

在后向传播时，输出层的偏差信号用计算机计算，它也可能包含导函数. 偏差信号的反向传播亦可用同样的前向型自适应 JTC 来完成. 修正后的互连可以通过 $LCTV_1$ 向系统输入新的 $g(x-a,y)$来建立. 用计算机对新的 $g(x-a,y)$进行计算的过程，其复杂系数的量级为 N^2 而不是 N^4. 而复杂系数为 N^4 的计算则由

JTC 实现.

8.4.3 Hopfield 模型

在上述学习方案中,总是假定学习过程是一个"试探-误差"操作即逐次逼近过程. 当学习的迭代过程结束后,互连网络就建立起来,识别操作只需一步就能完成. 然而,Hopfield 却设想识别过程可能不是一个一步就能完成的操作,而是一个逐步迭代的过程,在这一过程中输入信号是逐渐被识别的. 迭代过程一直到输入信号与内存中的参考信号达到最终匹配时才结束. 另一方面,学习过程,或互连的形成过程并非迭代的过程,而是遵循简单的法则.

在感知学习方法中,输入被分成两类. 在后向传播网络中,利用层层计算偏差信号,解决各类问题. 与它们不同,Hopfield 模型[59]注力于关联记忆的实现. 尽管这一模型也可以加以扩展,解决优化的问题[60]. 与其他神经网络一样,Hopfield 模型同样定义

$$z_j=1, \quad 若\sum_i w_{ji}x_i \geqslant T, \tag{18}$$

以及

$$z_j=0, \quad 若\sum_i w_{ji}x_i < T. \tag{19}$$

然而 z_j 并不是最终输出,它们被馈给系统作为新的输入 x_i,迭代过程一直到 z_j 与 x_i 相等时才结束. 在没有迭代的情况下,互连权重如下确定:

$$当\ j\neq i\ 时, \quad w_{ji}=\sum_s (2x_j^s-1)(2x_i^s-1), \tag{20}$$

$$当\ j=i\ 时, \quad w_{ji}=0, \tag{21}$$

其中 s 表示记忆的个数. 在 Hopfield 模型中,已知输入的出现将导致一个与该输入相同的输出,因为输入是事先存储在矩阵或张量存贮器中的. 尽管输入可能并不完备或有缺陷,它毕竟能通过迭代使网络完整地回忆起来. Hopfield 模型实际上是自动关联记忆.

Hopfield 模型是用光学方法实现的第一个神经网络[1]. 以前讨论过的方法当然可以用来实现(18)、(19)式中的张量-矩阵运算. 在大多数方案中,用电子学方法来完成(18)、(19)式中取阈值运算及(20)式中的互连表达式的建立. Jutamulia 等[19,61]认为互连表达式可以借助于液晶 SLM 及光致发光薄膜来建立,因为(20)式实质上仍然是线性运算.

当存储的向量或矩阵 x^i 正交时,Hopfield 模型十分适用;如果它们不正交时,在误差变得严重以前,仅有 0.15 N 个状态得以同时回忆起来,其中 N 是向量或矩阵的阶. 这个问题可以通过正交化手续加以克服,使系统性能有所改善,例如采用 Lu 等提出来的图形间联想方案(interpattern association,IPA)[62].

8.4.4 光学关联记忆

我们看到,Hopfield 模型确实就是一种自动关联记忆或联想记忆. 当我们想用光学方法实现 Hopfield 模型时,我们就完完全全回到了光学中,因为从本质上讲光学是能够产生关联记忆的,例如,全息就是典型的关联记忆[63~66]. 在重建时,用原来的光束之一去照射该全息图,就重建了原来的另一光束. 倘若两个光束来自相同的图形,全息术就是一种关联记忆. 大家知道,即使用一个不完全的图形(或有缺损的图形),都可以重构完整的图形. 输入图形缺损的程度不过是影响了信噪比而已.

常规的全息关联记忆和神经网络型的关联记忆的主要差别如下:全息关联记忆是一步到位的,它的信噪比取决于输入图形的缺损程度. 而神经网络型关联记忆是一个包含非线性运算的迭代过程,回忆起来的图形的信噪比在这一迭代过程中逐步改善. Owechko 等[67], Yariv 等[69], Anderson 和 Erie[70], Paek 和 Psaltis[68]提出一种神经网络型全息关联记忆,其中包括非线性取阈值运算和迭代. 也应当注意到,Mada 提议把经典的全息关联记忆作为一个普通的元件,用于构造一个光学计算机[71]. Yu 和 Ju-

tamulia 则阐明了光学关联记忆在符号替换运算中的应用[72].

全息关联记忆数学上可用外积来描述,然而外积可以转化为内积的形式,这样一来,关联记忆就可以借助于各种能进行内积运算的光学系统来实现[24,68,73,74].

8.4.5 无监视的神经网络

上面讨论过的那种通过一对输入-输出信号,去确定神经网络的互连权重因子的过程,称为监视下的学习或有监督的学习,其中的网络可以带隐蔽层,也可以不带隐蔽层. 如果不存在已知的输入-输出对,或者有一系列输入可以得到,我们只能循着一个组织起来的顺序去记忆它们,这一学习过程称无监视的学习. 无监视神经网络的例子是 Kohonen 自组织映射(KSOM)[49,75]. Lu 等描述了一个用光学方法实现 KSOM 的系统,在其中应用了透镜列阵[75].

为了实现 KSOM,图 8.20 中输入和互连权重都得由 $LCTV_1$ 显示,图中 $LCTV_2$ 可以移去,因为 KSOM 只需一个 LCTV,与 JTC 的原设计相同[26],然而有个参考函数. 为简单起见,仅讨论二维互连中的一维列阵,它们是 $g_1(x-a_1, y), g_2(x-a_2, y), \cdots, g_i(x-a_i, y), \cdots$. 与反向传播监视型网络不同,$g_i(x-a_i, y)$ 表示一个二维互连权重,而不是一个映射的输出. 在开始时互连权重 $g_i(x-a_i, y)$ 是随机的;当第一个输入信号 $f_1(x+a, y)$ 呈现在 $LCTV_1$ 时,在 $(a+a_i, 0)$ 处建立起 f_i 和 g_i 的相关输出,称输出节点(output nodes). 用 CCD 探测位于 $(a+a_i, 0)$ 处的相关峰,并用计算机分析和选择最大相关峰,对于在第 i 个最大节点近傍提供最大相关输出的节点,经刷新的计算机输出互连权重为

$$g_i(x-a, y) = \alpha_i f_i(x+a, y) + (1-\alpha_i) g_i'(x-a_i, y), \quad (22)$$

其中 $g'_i(x-a_i, y)$ 原本是 $g_i(x-a_i, y)$,α_i 是压缩系数,在 0 和 1 之间. 当输出节点远离所选中的第 i 个最大节点时 α_i 是较小的. 事实上,当输出节点相距所选中的节点过远时,α_i 为 0. 其一般原则,

就是使那些远离所选中节点的输出节点对应的互连权重在学习过程中不受影响. 如果 $\alpha_{i,\max}=1$,则学习过程一步到位. $\alpha_{i,\max}<1$ 时学习以迭代式循序渐进. 如果第一个图形是完整的,则经过无监视的学习后,当第一个图形再度出现在网络输入端,所选中的输出节点($a+a_{i,\max}$,0)处的相关与其他节点相比是相当高的,这样输入图形就被确认了. 当第二个图形输入系统时,学习过程重新开始. 如第二个图形与第一个类似,则输出节点将很靠近第一个节点,它将给出很大的相关输出,并被选中. 另一方面,若第二个图形与第一个图形差别很大,虽然其输出节点远离第一个选中的节点,但仍然提供了极大值,在这种情况下第二个图形也将被选中. 这样一来,不仅输入图形本身被识别,相像的图形也将被记忆下来并组织成一个族.

8.5 结　论

本章主要讨论了实现混合型神经网络的光学系统,这些神经网络以互连和非线性运算为其特色. 光学系统就其本质而言,原则上是适合于实现互连操作的. 其中的非线性运算则可以用电子学和数字计算机方法来完成.

我们介绍了神经网络理论的基础,讨论了三种监视型、一种无监视型学习方法. 显然还有许多监视型和无监视型学习方法,限于篇幅不能一一介绍. 尽管我们集中讨论了可以实现混合神经网络的光学方案,其中主要是进行矩阵-向量乘法和张量-矩阵乘法运算,对全息关联记忆也作了简要的介绍.

本章的目的并不是介绍所有的混合光学神经网络的系统,而是着重阐明用光学系统实现神经网络的基本方法. 我们也没有详细地分析各种光学系统技术方面的特点,因为所有这些技术都仍处于研究阶段,因而对有关的性能和限制也没有加以细致的评价. 事实上,混合光学神经网络的进展取决于两个因素:(1) 神经网络

比常规计算优越之处；(2) 实现神经网络的光学技术比电子技术优越之处.

出人意料的是，神经网络在商业上的应用已在各个领域出现，诸如通信、信用卡防伪、手写体识别、金融预测、半导体加工控制、语音识别、频谱分类及医学应用等等，尽管这些网络是用电子学方法实现的，且大部分依赖于计算机软件. 混合光学神经网络的商品化尚不到时候，除非网络中的神经元需要 1000 个以上的互连通道去解决有关的课题. 有兴趣的读者，可以阅读 Jutamulia 编辑的有关光学神经网络的论文集[77].

参 考 文 献

[1] D. Psaltis, and N. Farhat, Optical information processing based on associative-memory model of neural nets with thersholding and feedback, *Opt. Lett.*, **10** (1985), 98-100.

[2] H. J. Caulfield, J. Kinser, and S. K. Rogers, Optical neural nerworks, *Proc. IEEE*, **77** (1989), 1573-1583.

[3] F. T. S. Yu, Optical neural networks: architecture, design and models. In Progress in Optics, Vol. 32 (Ed. E. Wolf) North-Holland Physics Publishing Co. (1993), 61-44.

[4] S. Jutamulia, and F. T. S. Yu, Overview of hybrid optical neural networks, *Opt. Laser Tech.*, **28** (1996), 59-72.

[5] F. T. S. Yu, and S. Jutamulia, Optical Signal Processing, Computing, and Neural Networks, Wiley, New York (1992).

[6] W. S. McCulloch, and W. A. Pitts, A logical calculus of the ideas immanent in nervous activity, *Bull Math. Biophys*, **5** (1943).

[7] C. F. Stevens, The neuron, Scientific American 241 (3) (September 1979), 54-65.

[8] N. H. Farhat, Neural net models and optical comptuing: a brief overview. In Ooptical and Hybrid Computing (Ed. H. H. Szu) *Proc.*, **SPIE 634** (1986), 307-311.

[9] J. W. Goodman, A. R. Dias, and L. M. woody, Fully parallel, high-speed incoherent optical method for performing discrete Fourier trandform, *Opt. Lett.*, **2** (1978), 1-3.

[10] M. Kranzdorf, B. J. Bigner, L. Zhang, and K. M. Johnson, Optical connectionist machine with polarizat-ion-based bipolar weight values, *Opt. Eng.*, **28** (1989), 844-848.

[11] I. Shariv, O. Gila, and A. A. Friesem, All-optical bipolar neural network with polarization- modulating neurons, *Opt. Lett.*, **16**(1991), 1692-1694.

[12] F. T. S. Yu, X. Yang, and T. Lu, Space time-sharing optical neural networks, *Opt. Lett.*, **16** (1991), 247-249.

[13] J-S. Jang, S-W. Jung, S-Y. Lee, and S-Y. Shin, Optical implementation of the Hopfield model for two-dimensional associative memory, *Opt. Lett.*, **13** (1988), 248-150.

[14] X-M. Wang, and G-G. Mu, Optical neural network with bipolar neural states, *Appl. Opt.*, **31** (1992), 4712-4719.

[15] C-M. Uang, G. Lu, and F. T. S. Yu, Unipolar interpattern association neural network, *Opt. Lett.*, **19** (1994), 52-54.

[16] G. R. Little, S. C. Gustafson, and R. A. Senn, Generalization of the backpropagation neural network learning algorithm to permit complex weight, *Appl. Opt.*, **29** (1990), 1591-1592.

[17] N. H. Farhat, and D. Psaltis, Optical implementation of associative memory based on models of neural networks. In Optical Signal Processing (Ed. J. L. Horner)Academic Press, San Diego (1987), 129-162.

[18] R. A. Athale, and C. W. Stirk, Compact architectures for adaptive neural nets, *Opt. Eng.*, **28** (1989), 447-455.

[19] S. Jutamulia, G. M. Storti, W. Seiderman, and J. Lindmayer, Compact vector-matrix multiplier system employing electron trapping materials, US Patent 5, 134, 686(1992).

[20] Y. Nitta, J. Ohta, S. Tai, and K. Kyuma, Optical learning neurochip with internal analog memory, *Appl. Opt.*, **32** (1993), 1264-1274.

[21] S. Wu, T. Lu, X. Yu, and F. T. S. Yu, An adaptive optical neural network using a high resolution vide monitor, Microwave & *Opt. Tech. Lett.*, **2** (1989), 252-257.

[22] S. Lin, L. Liu, and Z. Wang, Optical implementation of the 2-D Hopfield model for a 2-D associative memory, *Opt. Commun.*, **70** (1989), 87-91.

[23] J. Widjaja, J. Uozumi, and T. Asakura, Optical implementation of artificial synapses by means of the Wigner distribution approach, *Opt. Eng.*, **33** (1994), 1655-1658.

[24] S. Jutamulia, G. M. Storti, J. Limdmayer, and W. Seiderman, Use of electron trapping materials in optical signal processing. 2: two-dimensional associative memory, *Appl. Opt.*, **30** (1991), 2879-2884.

[25] P. Lalanne, P. Chavel, and J. Taboury, Optical inner-product implementation of meural networks model, *Appl. Opt.*, **28** (1989), 377-385.

[26] S. Jutamulia, Uses of joint transform correlators in supervised and unsupervised hybrid computational-optical neural networks, *Opt. Rev.*, **1** (1995), 39-40.

[27] C. S. Weaver, and J. W. Goodman, A technique for optically convolving two functions, *Appl. Opt.*, **5** (1966), 1248-1249.

[28] J. E. Rau. Detection of differences in real distributions, *J Opt Soc Am.* **56** (1966), 1490-1494.

[29] R. N. Bracewell, The Fourier Transforn and Its Application, 2nd Edn, McGraw-Hill, New York (1978).

[30] A. Vander Lugt, Sigmal detection by complex spatial filtering, *IEEE trans. Inf. Theory*, **IT-10** (1964), 139-145.

[31] T-H. Chao, and W. W, Stoner, Optical implementation of a featurebased neural network with application to automatic target recognition, *Appl. Opt.*, **32** (1993), 1359-1369.

[32] D. Psaltis, C. H. Park, and J. Hong, Higher order associative memories and their optical implementations. Neural Networks., **1** (1988), 149-163.

[33] J. Hong, S. Campbell, and P. Yeh, Optical pattern classifier with perceptron learning, *Appl. Opt.*, **29** (1990), 3019-3025.

[34] J. W. Goodman, Introduction to Fourier Optics. McGraw-Hill, New York (1968).

[35] S. Jutamulia, and T. Asakura, A general approach to noncoherent correlations, *J. Optics* (Paris), **16** (1985), 121-125.

[36] S. Jutamulia (Ed), Selected Papers on Optical Correlators, SPIE Milestone Series MS76 (1993).

[37] H. J. White, and W. A. Wright, Holographic implememtation of a Hopfield model with discrete weights, *Appl. Opt.*, **27** (1988), 331-338.

[38] V. A. Ivanov, B. S. Kiselyov, A. L. Mikaelian, and D. E. Okonov, Optoelectronic neuroprocessor based on holographic disk memory using 1-D hologram recording, *Opt. Mem. Neur Netw.* **1** (1992), 55-62.

[39] D. Psaltis, D. Brady, X-G. Gu, and S. Lin, Holography in artificial neural networks, Nature, 343 (January 25, 1990), 325-330.

[40] K. Wagner, and D. Psaltis, Multilayer optical learning networks, *Appl. Opt.*, **26** (1987), 5061-5076.

[41] L. Solymar, and D. J. Cooke, Volume Holography and Volume Gratings, Academic Press, New York (1981).

[42] R. Arrathoon, Historical perspective: optical crossbars and optical computing. In Digital Optical Computing (Ed. R. Arrathoon), *Proc. SPIE*, **752** (1987), 2-11.

[43] A. P. Goutzoulis, Comparison of laser diode and electronic residue position-coded look-up tables, *Opt. Eng.*, **28** (1989), 373-379.

[44] F. Ito, and K. Kitayama, Optical implementation of the Hopfield neural network using multiple fiber nets, *Appl. Opt.*, **28** (1989), 4176-4181.

[45] J. W. Goodman, F. J. Leonberger, S-Y. Kung, and R. A. Athale, Optical interconnections for VLSI systems, *Proc. IEEE*, **72** (1984), 850-866.

[46] R. A. Athale, Optical matrix algebraic processors: a survey, Proc 10th Int. Opt. Comp. Conf. IEEE Cat. No 83CH 1880-4 (1983), 24-31.

[47] A. A. Sawchuck, B. K. Jenkins, C. S. Raghavendra, and A. Varma, Optical crossbar networks, Computer (June 1987), 50-60.

[48] F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain, *Psych. Rev.*, **65** (1958), 368-408.

[49] R. P. Lipmann, An introduction to computing with neural nets, *IEEE ASSP Magazine* (April 1987), 4-22.

[50] B. Widrow, and M. E. Hoff, Adaptive switching circuits, IRE WESCON Conv. Record Part 4 (1960), 96-104.

[51] A. D. McAulay, J. Wang, and X. Xu, Optical perceptron learning for binary classification with spatial light rebroadcasters, *Appl. Opt.*, **32** (1993), 1346-1353.

[52] F. Itoh, K. Kitayama, and Y. Tamura, Optical outer-product learning in a neural network using optically stimulable phosphors, *Opt. Lett.*, **15** (1990), 860-862.

[53] A. D. Fisher, W. L. Lippincott, and J. N. Lee, Optical implementations of associative networks with versatile adaptive learning capabilities, *Appl. Opt.*, **26** (1987), 5039-5054.

[54] D. E. Remelhart, G. E. Hinton, and R. J. Williams, Learning internal representations by erorr propagation. In Parallel Distributed Processing: Vol. 1 (Eds D. E. Rumelhart and J. L. McClelland,) MIT Press, Cambridge (1998), 318-364.

[55] S. Jutamulia, An implementation of backpropagation learning using an adaptive joint transform correlator, *Opt. Mem. Neur. Netw.*, **2** (1993), 177-183.

[56] C. Peterson, S. Redfield, J. D. Keeler, and E. Hartman, Optoelectronic implementation of multilayer neural networks in a single photorefractive crystal, *Opt. Eng.*, **29** (1990), 359-368.

[57] D. Psaltis, D. Brady, and K. Wagner, Adaptive optical networks using photorefractive crystals, *Appl. Opt.*, **27** (1988), 1752-1759.

[58] D. Kagan, and H. Friedman, Backpropagating neurons from bichromatic interaction with a three-level system, *Appl. Opt.*, **28** (1989), 1697-1700.

[59] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. of the National Academy of Sciences (USA), **79** (1982), 2554-2558.

[60] J. J. Hopfield, and D. W. Tank, Computing with neural circuits: a model, *Science*, **233** (1986), 625-633.

[61] S. Jutamulia, G. M. Storti, J. Lindmayer, and W. Seiderman, Use of electron trapping materials in optical signal processing. 3: modifiable Hopfield type neural networks, *Appl. Opt.*, **30** (1991), 1786-1790.

[62] T. Lu, X. Xu, S. Wu, and F. T. S. Yu, Neural network model using interpattern association, *Appl. Opt.*, **29** (1990), 284-288.

[63] P. J. van Heerden, A new optical method of storing and retrieving information, *Appl. Opt.*, **2** (1963), 387-392.

[64] P. J. van Heerden, Theory of optical information storage in solids, *Appl. Opt.*, **2** (1963), 393-400.

[65] D. Gabor, Associative holographic memory, *IBM J. Res. Dev.*, **13** (1969), 156-159.

[66] A. W. Lohmann, and Ch. Thum, Two-way code translation by computer-generated holographic filters, *Opt. Commun.*, **46** (1983), 74-78.

[67] Y. Owechko, G. J. Dunning, E. Marom, and B. H. Soffer, Holographic associative memory with nonlinearities in the correlation domain, *Appl. Opt.*, **26** (1987), 1900-1910.

[68] E. G. Paek, and D. Psaltis, Optical associative memory using Fourier transform holograms, *Opt. Eng.*, **26** (1987), 428-433.

[69] A. Yariv, S-K. Kwong, and K. Kyuma, Demonstration of an all-optical associative memory, *Appl. Phys. Lett.*, **48** (1986), 1114-1116.

[70] D. Z. Anderson, and M. C. Erie, Resonator memories and optical novelty filters, *Opt. Eng.*, **26** (1987), 434-444.

[71] H. Mada, Architecture for optical computing using holographic associative memories, *Appl. Opt.*, **24** (1985), 2063-2066.

[72] F. T. S. Yu, and S. Jutamulia, Implementation of symbolic substitution logic optical associative memories, *Appl. Opt.*, **26** (1987), 2293-2294.

[73] R. A. Athale, H. H. Szu, and C. B. Friedlander, Optical implementation of associative memory with controlled nonlinearity in the correlation domain, *Opt. Lett.*, **11** (1986), 484.

[74] H-K. Liu, S. Y. Kung, and J. A. Davis, Real-time optical associative retrieval technique, *Opt. Eng.*, **25**(1986), 853-856

[75] T. Kohonen, Self-organized formation of topologicallg correct feature maps, *Biological Cybernetics*, **43** (1982), 59-69.

[76] T. Lu, F. T. S. Yu, and D. A. Gregory, Self-organizing optical neural network for unsupervised learning, *Opt. Eng.*, **29** (1990), 1107-1113.

[77] S. Jutamulia(Ed). Seleeted papers on optical neural networks, SPIE milestone series Ms 96 (1994).

第九章　光折变介质信息处理

9.1　相位共轭和非线性介质中的相位共轭效应

9.1.1　相位共轭

如果光波 $E_1(\boldsymbol{r},t)$经过一个器件后产生一个反射波 $E_2(\boldsymbol{r},t)$，而 $E_2(\boldsymbol{r},t)$与 $E_1(\boldsymbol{r},t)$的波前完全重合，只是传播方向相反，那么 $E_2(\boldsymbol{r},t)$称为 $E_1(\boldsymbol{r},t)$的相位共轭波，能够产生相位共轭波的器件称为相位共轭器件.

相位共轭器件与常规的反射镜有本质上的区别，如图 9.1 所示. 当平面波 $\boldsymbol{E}_1$ 倾斜入射到反射镜上时，它将遵循反射定律，反射波 $\boldsymbol{E}_2$ 和入射波 $\boldsymbol{E}_1$ 关于法线 $\boldsymbol{n}$ 对称分布，$\boldsymbol{E}_2$ 和 $\boldsymbol{E}_1$ 在反射镜表面的切向分量相等.



(a) 光波在反射镜表面反射　　(b) 相位共轭镜

图 9.1　相位共轭镜与反射镜

然而当平面波 $\boldsymbol{E}_1$ 倾斜入射到相位共轭镜上时，情况就完全不同. 反射波 $\boldsymbol{E}_2$ 将按原路折回. 因此，当一个发散球面波射入相位共轭器件时，反射波将是一个会聚的球面波，其波前与入射波处处重合，但它们的幅度可能差一个常数因子. 特别是当光波通过一个相位畸变器件，再经过相位共轭镜反射，当反射波再次通过该相位畸变器件后，仍然恢复成原来的光波，这一效应称为相位恢复.

设标量波 E_1 沿 $+z$ 方向传播，它是一个单色波，但不一定是平面波，则 E_1 可表为

$$E_1(\boldsymbol{r},t)=\psi(\boldsymbol{r})\mathrm{e}^{\mathrm{i}(kz-\omega t)}=A_1(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\omega t},\tag{1}$$

式中

$$A_1(\boldsymbol{r})=\psi(\boldsymbol{r})\mathrm{e}^{\mathrm{i}kz}.\tag{2}$$

如果我们用某一方法产生另一光波 E_2，表为

$$E_2(\boldsymbol{r},t)=A_2(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\omega t},\tag{3}$$

且满足条件

$$A_2(\boldsymbol{r})=\alpha A_1^*(\boldsymbol{r})=\alpha\psi^*(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}kz},\tag{4}$$

式中 α 是常数，则称 E_2 是 E_1 的相位共轭波. 由上式可知 E_2 是沿 $-z$ 方向传播的波，E_2 和 E_1 具有相同的时间变化规律. 下面我们来证明当 E_2 满足条件(4)时，它的空间分布与 E_1 一致.

E_1 应满足亥姆霍兹方程：

$$\nabla^2E_1-\mu\varepsilon\frac{\partial^2E_1}{\partial t^2}=0,\tag{5}$$

其中 ε 为介电常数，μ 是导磁系数. 在定频的情况下，上式化作

$$\nabla^2E_1+\omega^2\mu\varepsilon E_1=0.\tag{6}$$

将(1)式代入(6)式，假定介电常数 ε 为实的，且在空间的变化足够缓慢，则有

$$\nabla^2\psi_1-[\omega^2\mu\varepsilon-k^2]\psi_1-2\mathrm{i}k\frac{\partial\psi_1}{\partial z}=0.\tag{7}$$

取上式的复共轭，得到

$$\nabla^2\psi_1^*+[\omega^2\mu\varepsilon-k^2]\psi_1^*+2\mathrm{i}k\frac{\partial\psi_1^*}{\partial z}=0,\tag{8}$$

上式的解是沿 $-z$ 方向传播的波，具有形式

$$C\psi_1^*(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}(kz+\omega t)}=A_2(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\omega t},\tag{9}$$

式中 C 是常数. 与(3)式相比较，发现它正是 E_2. 由于 E_2 和 E_1 满足同一个波动方程，根据麦克斯韦方程组作为电磁场运动方程的

完整性[4]，我们可以得到结论：E_2 和 E_1 的空间复振幅分布必然相同（可以差一个常数因子）.

9.1.2 由四波混频效应产生相位共轭

本节中我们将讨论非线性频率变换效应，亦即在电磁波传播时，由于介质的非线性极化产生的电磁波，它的频率、或偏振方向、或传播方向与入射波不同.

设有同一频率的两束光 $\boldsymbol{E}_1$ 和 $\boldsymbol{E}_4$，以不同方向射入介质（图 9.2），它们相干叠加的结果，必然会在介质中形成光强的周期分布，即光栅，可以表为

$$I(\boldsymbol{r})=I_0(1+m\cos(\boldsymbol{K}\cdot\boldsymbol{r})), \tag{10}$$

式中

$$I_0=I_1+I_4, \tag{11}$$

I_1 和 I_4 分别是两束入射光的光强，

$$m=2\sqrt{I_1 I_4}/I_0, \tag{12}$$

$$\boldsymbol{K}=\boldsymbol{k}_4-\boldsymbol{k}_1, \tag{13}$$

$$K=2k_1\sin\theta, \tag{14}$$

θ 是 $\boldsymbol{k}_1$ 和 $\boldsymbol{k}_4$ 的夹角的一半，参见图 9.3. $\boldsymbol{K}$ 与光栅方向正交，称为光栅的波矢量（注意 $\boldsymbol{k}$ 是波矢，不要混淆），且有

$$K=2\pi/\Lambda, \tag{15}$$

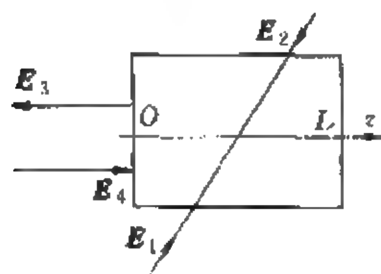


图 9.2 四波混频

E_1、E_2 为泵浦光，E_4 为探测光

E_3 为相位共轭光

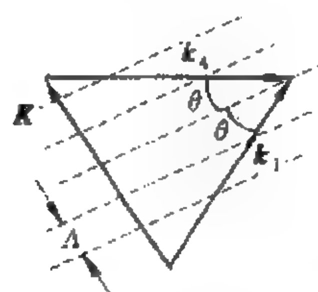


图 9.3 光栅波矢量 $\boldsymbol{K}$ 与波矢 $\boldsymbol{k}_1$、$\boldsymbol{k}_4$ 的关系

式中 Λ 为光栅的空间周期. 亦即同频率的两束光相干叠加的结果,在介质中产生了一个稳定的电场,强度呈周期性分布. 介质将在外电场的作用下发生极化,极化的结果在介质表面及内部出现周期性的束缚电荷分布. 束缚电荷激发起新的电场,叠加到由光波形成的电场上. 极化强度的分布决定了束缚电荷的分布,反过来它又为束缚电荷所制约,总电场 $\boldsymbol{E}$、极化强度矢量 $\boldsymbol{P}$ 和束缚电荷 ρ 的稳定分布是一个错综复杂、相互制约的过程,最后达到平衡态.

极化强度的周期性分布,相当于在介质中存在一个相位光栅,极化强度的周期性变化在某些介质中将引起折射率的周期性变化,类似于经过漂白的全息光栅.

当介质中存在第三束光 $\boldsymbol{E}_2$ 时,一般来讲三束光之间将两两形成光栅,情况更为复杂. 但我们总可以改变某些条件,例如偏振态、相干性等,使其中两束光,例如 $\boldsymbol{E}_1$ 和 $\boldsymbol{E}_4$ 形成的光栅起决定性作用. 这样一来,$\boldsymbol{E}_2$ 将在光栅中产生衍射. 当条件

$$\boldsymbol{k}_2 = -\boldsymbol{k}_1 \tag{16}$$

满足时,即 $\boldsymbol{E}_2$ 和 $\boldsymbol{E}_1$ 反向传播时,它恰恰满足 Bragg 衍射的条件,

$$2k_2\sin\theta_2 = 2k_1\sin\theta = 2\pi/\Lambda. \tag{17}$$

此时出现共振现象:全部(或大部)衍射光都集中在一级衍射波内,称为 Bragg 衍射,且衍射光 $\boldsymbol{E}_3$ 与 $\boldsymbol{E}_4$ 反向传播,成为 $\boldsymbol{E}_4$ 的相位共轭波. 这就是四波混频产生相位共轭光波的简单物理图象. $\boldsymbol{E}_1$ 和 $\boldsymbol{E}_2$ 通常称为泵浦光,$\boldsymbol{E}_4$ 称为探测光(或信号光),$\boldsymbol{E}_3$ 称为相位共轭光,见图 9.4. 有时称 $\boldsymbol{E}_1$ 和 $\boldsymbol{E}_4$ 为写入光,因为它们的相干叠加在介质中形成体光栅. $\boldsymbol{E}_2$ 称为读出光,$\boldsymbol{E}_3$ 则为相位共轭光. 当读出光 $\boldsymbol{E}_2$ 与 $\boldsymbol{E}_1$,$\boldsymbol{E}_4$ 写入的光栅的夹角 $\theta_2=\theta$ 时,满足 Bragg 条件,产生相位共轭光波 $\boldsymbol{E}_3$.

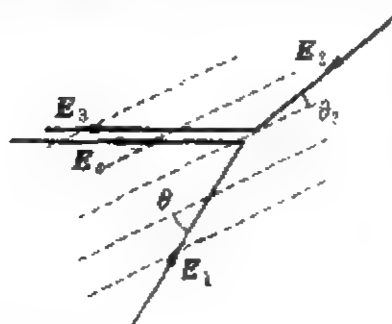


图 9.4 四波混频的简单物理机制

虽然 E_4 的入射方向是任意的，然而只要条件(16)满足，由读出光 E_2 形成的衍射光 E_3 一定和 E_4 反向传播．这正是相位共轭的基本要求．

以上讨论中，我们实际上假设所用的光波都是平面波．下面我们将证明这一假设并不必要．

9.1.3 利用非线性介质中的四波混频效应产生相位共轭[6]

在图 9.2 所示的四波混频机构中，设

$$\left.\begin{aligned} E_1 &= A_1(\boldsymbol{r})\mathrm{e}^{[\mathrm{i}(\boldsymbol{k}_1\cdot\boldsymbol{r}-\omega t)]}, \\ E_2 &= A_2(\boldsymbol{r})\mathrm{e}^{[\mathrm{i}(\boldsymbol{k}_2\cdot\boldsymbol{r}-\omega t)]}, \\ E_4 &= A_4(\boldsymbol{r})\mathrm{e}^{[\mathrm{i}(\boldsymbol{k}_4\cdot\boldsymbol{r}-\omega t)]}, \end{aligned}\right\} \tag{18}$$

这里假设三个波的偏振方向相同，因此我们可以用标量来表示它们．由于 A_1，A_2 和 A_4 都是 $\boldsymbol{r}$ 的函数，所以它们并不一定是平面波．这三个光波在介质中产生的极化强度可表为

$$P' = \frac{1}{2}\chi^{(3)}A_1A_2A_4^*\,\mathrm{e}^{\{\mathrm{i}[(\boldsymbol{k}_1+\boldsymbol{k}_2-\boldsymbol{k}_4)\cdot\boldsymbol{r}-(\omega_1+\omega_2-\omega_4)t]\}}, \tag{19}$$

式中 $\chi^{(3)}$ 为极化率的三阶分量．设 $\boldsymbol{e}_z$ 为 z 方向的单位矢量，则在条件

$$\left.\begin{aligned} \boldsymbol{k}_2 &= -\boldsymbol{k}_1, \\ \boldsymbol{k}_4 &= k\boldsymbol{e}_z, \\ \omega_1 &= \omega_2 = \omega_4 = \omega \end{aligned}\right\} \tag{20}$$

满足时，(19)式化为

$$P' = \frac{1}{2}\chi^{(3)}A_1A_2A_4^*\,\mathrm{e}^{-\mathrm{i}(kz+\omega t)}, \tag{21}$$

该极化强度反过来又将激发出一个沿 $-z$ 方向传播的波，记为 E_3

$$E_3 = A_3(z)\mathrm{e}^{-\mathrm{i}(kz+\omega t)}, \tag{22}$$

而 E_3 和 E_1、E_2 耦合产生了新的极化，记为

$$P'' = \frac{1}{2}\chi^{(3)}A_1A_2A_3^*\,\mathrm{e}^{-\mathrm{i}(kz-\omega t)}. \tag{23}$$

注意 P''和 E_4 传播方向相同，频率相同，因此产生了强烈的耦合。我们可以把上述过程看成是 E_4，E_3 通过与泵浦光 E_1，E_2 的耦合在非线性介质中互相作用的结果。可以认为 E_3 的空间变化是 P'引起的，而 E_4 的空间变化则是 P''引起的，E_3，E_4 相对于泵浦光 E_1，E_2 是微扰，从而写出下面的耦合波方程：

$$\left.\begin{aligned} dA_3/dz &= i\kappa^* A_4^*, \\ dA_4^*/dz &= i\kappa A_3, \end{aligned}\right\} \tag{24}$$

式中

$$\kappa^* = \frac{\omega}{2}\sqrt{\frac{\mu_0}{\epsilon}}\chi^{(3)} A_1 A_2. \tag{25}$$

假定 E_3 和 E_4 对 E_1 和 E_2 没有影响，即设 A_1A_2 为常数，则(24)式的解为

$$\left.\begin{aligned} A_3(z) &= \frac{\cos(|\kappa|z)}{\cos(|\kappa|L)}A_3(L) + i\frac{\kappa^*\sin[|\kappa|(z-L)]}{|\kappa|\cos(|\kappa|L)}A_4^*(0), \\ A_4^*(z) &= i\frac{|\kappa|\sin(|\kappa|z)}{\kappa^*\cos(|\kappa|L)}A_3(L) + \frac{\cos[|\kappa|(z-L)]}{\cos(|\kappa|L)}A_4^*(0). \end{aligned}\right\} \tag{26}$$

设 $A_3(L)=0$，即 E_3 没有透射波，则

$$\left.\begin{aligned} A_3(z) &= i\frac{\kappa^*\sin[|\kappa|(z-L)]}{|\kappa|\cos(|\kappa|L)}A_4^*(0), \\ A_4^*(z) &= \frac{\cos[|\kappa|(z-L)]}{\cos(|\kappa|L)}A_4^*(0). \end{aligned}\right\} \tag{27}$$

再设 $z<0$ 的区域内(晶体外)E_4 具有下面的形式：

$$E_4'' = \psi(\boldsymbol{r})e^{i(kz-\omega t)}, \tag{28}$$

在 $0<z<L$(介质内)

$$E_4' = \frac{\cos[|\kappa|(z-L)]}{\cos(|\kappa|L)}A_4(0)e^{i(kz-\omega t)}, \tag{29}$$

在 $z=0$(界面)处的连接条件给出

$$\psi_0(\boldsymbol{r}) = \psi(\boldsymbol{r})|_{z=0} = A_4(0), \tag{30}$$

由(27)及(30)式，在介质内 E_3 可表为

$$E_3' = \mathrm{i}\,\frac{\kappa^*\sin[|\kappa|(z-L)]}{|\kappa|\cos(|\kappa|L)}A_4^*(0)\mathrm{e}^{-\mathrm{i}(kz+\omega t)}$$

$$= \mathrm{i}\,\frac{\kappa^*\sin[|\kappa|(z-L)]}{|\kappa|\cos(|\kappa|L)}\psi_0^*(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}(kz+\omega t)}, \tag{31}$$

在 $z<0$ 时，E_3 可取为

$$E_3'' = -\mathrm{i}\left[\frac{\kappa^*}{|\kappa|}\tan(|\kappa|L)\right]\psi^*(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}(kz+\omega t)}, \tag{32}$$

这样一来，E_3'' 表示一个向 $-z$ 方向传播的波，E_3'' 和 E_3' 在界面上满足连接条件，而且在介质外($z<0$)，总有

$$A_3(\boldsymbol{r}) = 常数 \times A_4^*(\boldsymbol{r}), \tag{33}$$

上式表明 E_3 在晶体外是 E_4 的相位共轭波.

9.2 光折变效应相位共轭器

9.2.1 引言

自从四波混频作为一个产生相位共轭的有效方法提出来以后，人们就开始寻找这样一类介质，在其中四波混频具有足够高的效率. 上一节介绍的一般非线性介质中，三阶非线性系数 $\chi^{(3)}$ 都非常小，要引起显著的效应，必须有足够高的输入光功率，例如 1 MW/cm^2 甚至更高. 然而，在一类称为光折变的晶体材料中，只要很小的输入光功率，例如 1 W/cm^2，就足以引起相当显著的效应，产生足够强的相位共轭光波. 光折变晶体种类很多，例如 $BaTiO_3$，$KTa_{1-x}Nb_xO_3$ (KTN)，$Sr_{1-x}Ba_xNb_2O_6$ (SBN)，$Ba_{2-x}Sr_xK_{1-y}Na_yNb_5O_{15}$ (BSKNN)，$Bi_{12}GeO_{20}$ (BGO)，$Bi_{12}SiO_{20}$ (BSO)，$LiNbO_3$，$KNbO_3$，GaAs，InP，CdTe 等.

光折变效应引起的四波混频-相位共轭效应可简单归纳为下面的过程：

(1) 输入光波(写入光)引起晶体材料中电荷的迁移和正、负电荷的分离. 特别是当两个光波在晶体中相干叠加形成光强的周

期性分布时，电荷分布也呈现周期性.

(2) 正、负电荷之间产生很强的内部静电场，达 10^5 V/m 的数量级以上.

(3) 光折变晶体都是电光晶体，静电场通过线性电-光效应(Pockels 效应，参见节 7.6.1)，引起晶体中折射率的周期性分布，它相当于相位光栅或体全息图.

(4) 当第三个光波(读出光)射入晶体，且满足 Bragg 条件时，它被相位光栅衍射，衍射光正是某一写入光的相位共轭光波. 严格来讲，相位共轭光波是上述三个光波共同作用产生的，而光折变晶体的非线性效应，是这些光波相互作用产生相位共轭波的根本原因.

光折变效应引起的相位共轭具有以下特点：

首先，光波之间的耦合，是通过光波在光折变介质中写入的相位光栅而间接实现的，而相位光栅的建立和擦除则是上述的电场诱导的电荷重新分布的过程，需要一定的时间，可称为光折变效应的特征时间或弛豫时间 τ. 相位共轭波是读出光被相位光栅衍射造成的，因而从时序上看，相位共轭波的建立和稳定要比写入光延迟一段时间 τ，所以引起某些特殊配置下瞬态响应特性和稳态响应特性的差异.

其次，在产生相位共轭波的过程中，必然伴随着能量在光束间的转移和重新分配. 由于许多光折变材料具有很高的线性电光系数，例如 $BaTiO_3$ 的 γ_{42}高达 1300 pm/V，使得很弱的泵浦光就可以导致很显著的折射率变化，形成相位光栅，光束通过相位光栅的耦合造成能量转移. 这种转移在某些特定的条件下可以达到衍射光把泵浦光耗尽的程度.

第三，光折变效应还具有积累效果，亦即很微弱的输入光经过充分长时间，可以形成很强的内电场和折射率变化. 这显然是电荷迁移和重新分布的积分效应.

这些显著的特点，形成了一系列复杂、特殊的效应，例如自泵

浦相位共轭(SPPC),互泵浦相位共轭(MPPC 或 DPC),光放大、光振荡等,成为实时光学信息处理的重要方法,广泛应用于光信息存储、图像相加、相减,像差图像复原、边缘突出、相关识别、光互连及光计算中,光折变已成为非线性光学中一个重要的效应. Golomb 等[1]、Yeh[2,3]和刘思敏等[5]对光折变效应的机理及应用进行了详细的讨论.

9.2.2 光折变效应

在光折变介质中,存在一定数量的束缚状态,由杂质或缺陷(如缺位、间隙原子、位错)引起. 电子可能由这些杂质或缺陷所束缚,束缚电子的能级称杂质能级,它位于导带和价带之间. 这类由杂质或缺陷提供的带有电子的能级称施主.

Yeh 提出一个简单的模型[3]. 设在光折变介质中所有的施主杂质都相同,并处于带隙的同一能级上,见图 9.5.

光折变介质受到光照时,杂质施主中的电子以一定的概率被激发,由杂质能级跃迁到导带中去,电离后的施主又可能俘获新的电子. 设施主的浓度为 N_D,其中 N_D^i 被电离,则产生电子的速率取决于未被激发的施主浓度,即

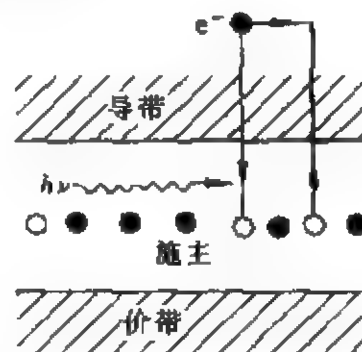


图 9.5 光折变介质的能带模型[3]

$$r_e=(\sigma I+\beta)(N_D-N_D^i), \quad (1)$$

式中 β 为热激发产生电子的速率,σI 则是光激发产生电子的速率,I 为光强,σ 为光激发截面. 设热激发的速率远小于光激发的速率,$\beta \ll \sigma I$,(1)式中热激发的贡献可略去,则在 Δt 内电离的施主的增量为

$$\Delta N_D^i=\sigma I(N_D-N_D^i)\Delta t-\gamma_R N N_D^i \Delta t, \quad (2)$$

式中第二项代表电离的施主重新俘获电子的变化量,它显然正比于电子的浓度 N 和电离的施主浓度 N_D^i,式中 γ_R 是俘获截面. 上

式可改写为

$$\frac{\partial N_D^i}{\partial t}=\sigma I(N_D-N_D^i)-\gamma_R N N_D^i. \tag{3}$$

考虑一个体积元 ΔV，在其中 Δt 内由光激发引起的导带电子的增量为 $N_D^i\Delta V$. ΔV 内总的电子增量则为 $N\Delta V$，它们的差显然是电子的迁移即位移电流造成的，因此我们有

$$-\oint\!\!\!\oint_{(S)} \boldsymbol{j}\cdot d\boldsymbol{S}=-q\frac{\partial(N-N_D^i)}{\partial t}\Delta V, \tag{4}$$

式中 $\boldsymbol{j}$ 是电流密度，$-q$ 为电子电量，左边的积分在 ΔV 的表面 S 上进行. 当 $\Delta V\to 0$ 时，我们得到速率方程

$$-q\left(\frac{\partial N_D^i}{\partial t}-\frac{\partial N}{\partial t}\right)=\nabla\cdot\boldsymbol{j}. \tag{5}$$

电荷载流子的移动引起电荷的重新分布，电荷的分布将导致电场 $\boldsymbol{E}$，该电场反过来又影响了载流子的位移. 这样一来，在介质中同时存在两个过程：一方面，载流子从浓度高的区域向浓度低的区域扩散；另一方面，载流子又在空间电场的作用下发生迁移，因此(5)式中的电流密度是两者共同的贡献：

$$\boldsymbol{j}=qN\mu\boldsymbol{E}+qD\,\nabla N, \tag{6}$$

式中 μ 是迁移率，D 是扩散系数，(6)式中前一项表征载流子在空间电场中的定向流动，后一项代表载流子浓度梯度导致的扩散. 当 $\boldsymbol{j}=0$ 时，达到平衡.

麦克斯韦方程（泊松方程）为

$$\nabla\cdot\varepsilon\boldsymbol{E}=4\pi\rho=-4\pi q(N-N_D^i), \tag{7}$$

式中 ρ 是电荷密度，ε 是介电张量.

在光折变介质受到光的辐照时，杂质施主可能吸收光子，释放电子而电离，电子跃迁到导带中，而电离后的施主又可能俘获导带中的电子. 光折变效应正是通过介质中的杂质施主在光照下的行为实现的. 下面我们考虑两束光照射到光折变介质中的情况.

设

$$E_1=A_1e^{i(k_1\cdot r-\omega t)}, \tag{8}$$

$$E_2=A_2e^{i(k_2\cdot r-\omega t)}, \tag{9}$$

我们设 $A_1=A_1(r)$是空间缓变的函数，而 A_2 为常数．两束光在介质中相干叠加形成的光强分布则为

$$I=I_1+I_2+A_1A_2^*e^{-iK\cdot r}+A_1^*A_2e^{iK\cdot r}=I_0[1+m\cos(\boldsymbol{K}\cdot\boldsymbol{r})], \tag{10}$$

式中

$$I_0=I_1+I_2, \tag{11}$$

$$m=2\sqrt{I_1I_2}/I_0, \tag{12}$$

$$\boldsymbol{K}=\boldsymbol{k}_2-\boldsymbol{k}_1, \tag{13}$$

$\boldsymbol{K}$ 的值

$$K=|\boldsymbol{K}|=2\pi/\Lambda, \tag{14}$$

Λ 表示光强分布的空间周期．

由于光强是周期性分布的，在亮区内，介质吸收光子产生的电子比暗区多，因而电子从亮区向暗区扩散，留下带正电荷的施主．这一扩散过程的结果，导致正电荷分布在亮区，负电荷分布在暗区．正负电荷的分离形成电场 $\boldsymbol{E}$，方向是从亮区指向暗区，它使扩散作用逐渐减缓，最后达到平衡．从而空间电荷密度 ρ 可表为

$$\rho=\rho_0\cos(\boldsymbol{K}\cdot\boldsymbol{r}), \tag{15}$$

ρ_0 是常数，注意 ρ 和光强 I 是同相的．以(15)式代入(5)式得到

$$\boldsymbol{E}=4\pi\rho_0\frac{\boldsymbol{K}}{\boldsymbol{K}\cdot\varepsilon\boldsymbol{K}}\sin(\boldsymbol{K}\cdot\boldsymbol{r}), \tag{16}$$

表明电场强度 $\boldsymbol{E}$ 与空间光强分布有 $\pi/2$ 的相位差，电场强度在电荷密度 ρ 的梯度最大处达到极值．

由于线性电光效应(Pockels 效应，参见节 7.6.1)

$$\Delta(1/n^2)=-2\Delta n/n^3=\gamma_{eff}E,$$

或

$$\Delta n=-\gamma_{eff}n^3E/2, \tag{17}$$

式中 γ_{eff}是在光折变效应中起主要作用的线性电光系数，它是线性

电光张量的一个分量，称为有效电光系数．在没有外加电场时，E 为空间电荷建立的电场．

Δn 与 E 可能反相也可能同相，取决于 γ_{eff} 的正负．在没有外加均匀电场时，Δn 与光强分布就有 $\pi/2$ 或 $-\pi/2$ 的相位差，这是光折变效应的重要特征．

光折变晶体中至少有一个线性电光系数较大，以引起足够明显的光折变效应．在 $BaTiO_3$ 中，γ_{42} 高达 1300 pm/V，因此 $BaTiO_3$ 是光折变效应极强的晶体．

图 9.6 给出光强 I，空间电荷密度 ρ，电场 E 和折射率改变量 Δn 的相位关系．图中，自上至下，I 为光波相干叠加形成光强的周期分布；ρ 是光致电荷密度分布，正电荷趋向于分布在光强的亮纹处；E 是空间电荷形成的电场强度，由泊松方程(7)决定，即电场强度的散度正比于电荷密度；Δn 为光折变介质中折射率的变化，它与 E 成比例．

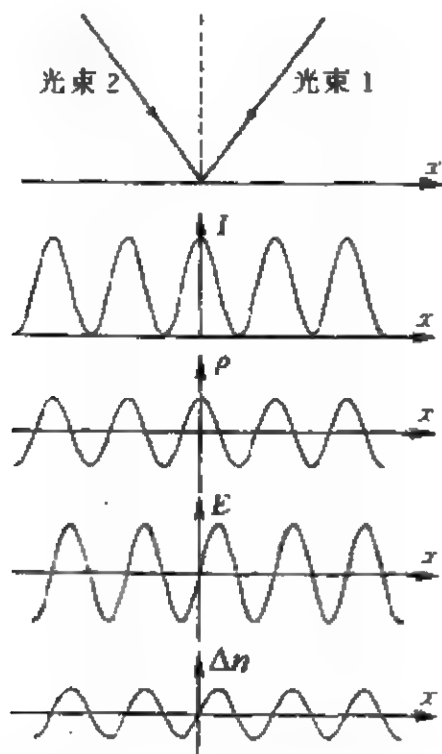


图 9.6　光折变介质中折射率光栅的形成[3]

9.2.3 全息光栅模型

参见图 9.7，设两个波 $\boldsymbol{E}_1$，$\boldsymbol{E}_4$ 具有相同的频率 ω. 为简单起见，设它们具有相同的偏振，因此我们可以用标量来处理. 设

$$E_1 = A_1 \exp[\mathrm{i}(\boldsymbol{k}_1 \cdot \boldsymbol{r} - \omega t)], \tag{18}$$

$$E_4 = A_4 \exp[\mathrm{i}(\boldsymbol{k}_4 \cdot \boldsymbol{r} - \omega t)], \tag{19}$$

其中 A_1 = 常数，即 E_1 是平面波，

$$A_4 = A_4(\boldsymbol{r}), \tag{20}$$

相对于迅变的相因子 $\exp[\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r})]$，$A_4(\boldsymbol{r})$ 是缓慢变化的调制包络，例如在该光波中放入一个透明片，上面有某种图案.

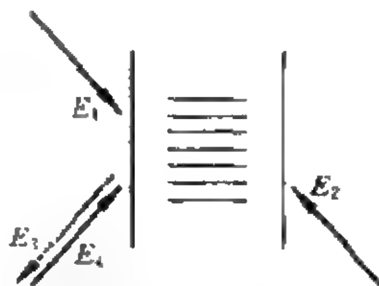


图 9.7 光折变晶体中的相位共轭效应

这两束光以一定的夹角射入光折变介质，并在介质中相干叠加形成条纹图，其强度分布为

$$I = |A_1|^2 + |A_4|^2 + A_1 A_4^* \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}} + A_1^* A_4 \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}, \tag{21}$$

式中

$$\boldsymbol{K} = \boldsymbol{k}_4 - \boldsymbol{k}_1. \tag{22}$$

这一强度分布在介质中引起光折变效应，Δn 可表为

$$\Delta n = \frac{n'}{2}\mathrm{e}^{\mathrm{i}\varphi'}\frac{A_1 A_4^*}{I_0}\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}} + \mathrm{c.c.}, \tag{23}$$

式中 c.c. 表示复数共轭，$I_0 = I_1 + I_4$. 上式表明，光折变效应引起的折射率变化与光强分布成比例，φ' 是 Δn 与 I 的相位差. n' 是一个适当的比例因子，它比未发生光折变效应的折射率 n 小一个量级以上.

晶体中的相位光栅 $n(\boldsymbol{r})$ 与全息光栅非常相似，因此它也会发生类似全息再现的效应，E_4 为物光，E_1 为参考光. 下面我们考虑一个简单的模型. 设相位光栅

$$n = n_0 + \Delta n(\boldsymbol{r}). \tag{24}$$

假设第三个光波（读出光）

$$E_2 = A_2 \exp[-i(\boldsymbol{k}_1 \cdot \boldsymbol{r} + \omega t)], \tag{25}$$

即该光波的传播方向与 E_1 相反，设该光波与相位光栅发生相互作用，使电位移矢量有一个改变量

$$\begin{aligned} \Delta D &= \Delta n^2 \cdot E_2 = 2n\Delta n E_2 \\ &= \left[nn' e^{i\varphi'} \frac{A_1 A_4^*}{I_0} e^{-i\boldsymbol{K}\cdot\boldsymbol{r}} \right] A_2 e^{-i\boldsymbol{k}_1\cdot\boldsymbol{r}} + \text{c.c.} \\ &= nn' e^{i\varphi'} \frac{A_1 A_4^* A_2}{I_0} e^{-i(\boldsymbol{k}_4\cdot\boldsymbol{r} - \omega t)} + \text{c.c.}, \end{aligned} \tag{26}$$

(26)式相当于三个波 A_1、A_4 和 A_2 相互作用产生了极化 P(参见9.1节(21)式)，它激发一个光波，记为

$$E_3 = Cn' e^{i\varphi'} (A_1 A_2) A_4^* \exp[-i(\boldsymbol{k}_4 \cdot \boldsymbol{r} + \omega t)], \tag{27}$$

这一光波的传播方向与 E_4 相反，其振幅是 A_4 的复共轭 A_4^*(差一个常数)，因此就是 E_4 的相位共轭波.

在这一模型中，对三个波之间的相互作用作了简单的处理，也不考虑它们之间的能量交换，可认为是零级近似. 在下一节中我们将用耦合波方法来处理光折变晶体中光波之间的相互作用，可认为是一级近似.

在光折变效应中，相位光栅起到了关键的作用，它是由物光和参考光的相干叠加形成的光强周期性分布、电荷的位移及线性电光效应的综合效果. 相位光栅又是物光空间信息的载体. 读出光只是读这一光栅，产生了物光的相位共轭光.

(23)式中的 n'，φ' 是表征光折变效应的两个主要参数. 对比(17)式及(23)式，知 n' 与光折变介质的有效电光系数 γ_{eff} 成正比.

以 SBN(铌酸锶钡)、$BaTiO_3$(钛酸钡)这一类 4 mm 族晶体为例，它们的电光系数张量可表为

$$\gamma = \begin{pmatrix} 0 & 0 & \gamma_{13} \\ 0 & 0 & \gamma_{13} \\ 0 & 0 & \gamma_{33} \\ 0 & \gamma_{42} & 0 \\ \gamma_{42} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{28}$$

当参与相互作用的光为寻常光时，

$$\gamma_{eff}=\gamma_{13}\sin\frac{\alpha+\beta}{2},\tag{29}$$

式中 α,β 分别是泵浦光 1,4 与晶体光轴的夹角. 当使用非常光时，

$$\gamma_{eff}=\frac{1}{n_e n_o}\Big[n_e^4\gamma_{33}\sin\alpha\sin\beta+2n_e^2n_o^2\gamma_{42}\cos^2\frac{\alpha+\beta}{2}+n_o^4\gamma_{13}\cos\alpha\cos\beta\Big]\sin\frac{\alpha+\beta}{2},\tag{30}$$

γ_{33}, γ_{42} 和 γ_{13} 是三个非零的线性电光系数，n_o 和 n_e 则是寻常光和非常光折射率.

上面已提到，$BaTiO_3$ 的三个 γ 中最大的是 γ_{42}，$\gamma_{42}=1300$ pm/V，而 SBN 的电光系数则为：$\gamma_{33}=244$ pm/V，$\gamma_{13}=55$ pm/V、$\gamma_{42}=80$ pm/V，可见在这两种晶体中引起光折变效应的都是非常光.

寻常光不参与光折变效应，常用来“擦除”由非常光产生的相位光栅.

9.3 用耦合波近似处理光折变晶体中的二波混频和四波混频效应

在光折变效应中，由于写入光波相干叠加形成的光强周期性分布与内电场通过电光效应形成的折射率光栅之间有相位差，因此，不能像节 9.1.3 中那样简单地用一个三阶非线性张量 $\chi^{(3)}$ 来处理.

节 9.2.3 的“全息光栅模型”假定两个写入光波首先在介质中形成折射率光栅，然后读出光波在该光栅上衍射产生相位共轭光. 而事实上包括相位共轭波在内的四个光波在介质中是同时存在的，它们通过介质的光折变效应互相耦合，彼此影响和制约，呈现出复杂的物理效应，仅仅用读出光在相位光栅上的衍射产生相位共轭光这样的模型来解释这一效应，又嫌过于粗糙.

本节中，我们采用缓变包络近似，用耦合波方法来处理光折变介质中光波的相互作用过程.

9.3.1 在缓变包络及单一光栅近似下介质中的波动方程

设介质中同时存在 N 个光波，它们的光矢量为 $E_j(j=1,2,\cdots,N)$，设它们的频率相同（均为 ω），偏振方向相同，因此可以用标量来表示：

$$E_j=A_j e^{i(\boldsymbol{k}_j\cdot\boldsymbol{r}-\omega t)}\quad(j=1,2,\cdots,N). \tag{1}$$

进一步假设这 N 个波都在 xz 平面内大致沿 z（或 $-z$）方向传播，且波场在 x 方向无限扩展，则边界条件要求 A_j 仅仅是 z 的函数.

这 N 个波在介质中通过光折变效应一般将会形成多个光栅. 为简单起见，我们假定其中一个光栅是主要的，该光栅的波矢量为 $\boldsymbol{K}$. 折射率的分布可表为

$$n=n_0+\Delta n, \tag{2}$$

式中 n_0 是没有光照下的折射率；

$$\Delta n=\frac{n'}{2}e^{i\varphi'}\frac{f(A_1,\cdots,A_N)}{I_0}e^{i(\boldsymbol{K}\cdot\boldsymbol{r})}+\text{c. c.}, \tag{3}$$

其中 c. c. 表示复数共轭；$f(A_1,\cdots,A_N)$ 一般来讲是各个波的振幅 A_j 的函数；I_0 为总光强，

$$I_0=I_1+I_2+\cdots+I_N; \tag{4}$$

n' 是正实数，比 n_0 小一个量级以上；φ' 是实数，表示相位光栅与光强周期分布之间的相位差. 由于 Δn 的存在，电位移矢量的分量

$$D=n^2E\approx(n_0^2+2n_0\Delta n)E, \tag{5}$$

其括号中第二项 $2n_0\Delta n$ 相当于介质中存在的非线性极化率 χ. 麦克斯韦方程组则为

$$\left.\begin{aligned}&\nabla\times\boldsymbol{H}=\frac{1}{c}\frac{\partial\boldsymbol{D}}{\partial t},\\&\nabla\times\boldsymbol{E}=-\frac{\mu}{c}\frac{\partial\boldsymbol{H}}{\partial t},\\&\nabla\cdot\boldsymbol{E}=0,\\&\nabla\cdot\boldsymbol{H}=0.\end{aligned}\right\}\tag{6}$$

由式(5)和(6)导出标量波动方程

$$(\nabla^2+n_0^2k_0^2)E+2n_0k_0^2\Delta nE=0,\tag{7}$$

式中 E 是各个光波的电场强度之和：

$$E=\sum_j E_j(\boldsymbol{r},t)=\sum_j\{A_j(z)\exp[\mathrm{i}(\boldsymbol{k}_j\cdot\boldsymbol{r}-\omega t)]+\mathrm{c.c.}\}.\tag{8}$$

将(8)式代入(7)式，设 A_j 是 z 的缓慢变化的函数，我们只保留 A_j 及它的一次导数，略去二次导数，得到

$$\begin{aligned}&\sum_j 2\mathrm{i}k_{jz}\left\{\frac{\mathrm{d}A_j}{\mathrm{d}z}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}_j\cdot\boldsymbol{r}-\omega t)}-\frac{\mathrm{d}A_j^*}{\mathrm{d}z}\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}_j\cdot\boldsymbol{r}-\omega t)}\right\}\\&\quad=-2n_0k_0^2\Delta n\sum_j\{A_j\mathrm{e}^{\mathrm{i}(\boldsymbol{k}_j\cdot\boldsymbol{r}-\omega t)}+\mathrm{c.c.}\}\\&\quad=-\frac{k_0^2n_0n'}{I_0}[f(A_1,\cdots,A_N)\mathrm{e}^{\mathrm{i}(\boldsymbol{K}\cdot\boldsymbol{r})+\varphi'}+\mathrm{c.c.}]\\&\qquad\times\sum_j\{A_j\mathrm{e}^{\mathrm{i}(\boldsymbol{k}_j\cdot\boldsymbol{r}-\omega t)}+\mathrm{c.c.}\}.\end{aligned}\tag{9}$$

给定各个波的波矢量 $\boldsymbol{k}_j$ 及函数 $f(A_1,\cdots,A_N)$ 的形式，方程(9)就可以求解. Golomb 等[1]和 Yeh[3]对二波混频及四波混频情况作了详细的讨论.

9.3.2 二波混频

参见图 9.8，设两个光波从晶体同侧($z=0$)射入晶体，它们关于 z 轴对称分布，在晶体中波矢量与 z 轴夹角为 θ，

$$k_{1z}=k_{2z}\approx n_0k_0\cos\theta.\tag{10}$$

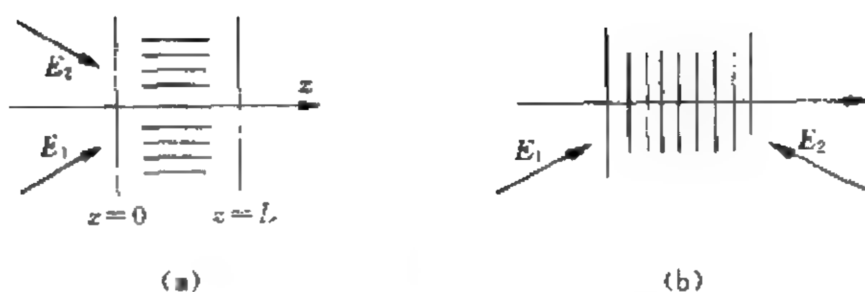


图 9.8 二波混频

(a) E_1, E_2 射入晶体的同一侧 (b) E_1, E_2 从晶体的两侧射入

在二波混频的情况下，由节 9.2(23)式，折射率变化引起的相位光栅可表为

$$\Delta n=\frac{n'}{2}e^{i\varphi'}\frac{A_1A_2^*}{I_0}e^{-i\boldsymbol{K}\cdot\boldsymbol{r}}+\text{c.c.}, \tag{11}$$

其中

$$I_0=I_1+I_2, \tag{12}$$

$$\boldsymbol{K}=\boldsymbol{k}_2-\boldsymbol{k}_1. \tag{13}$$

此时有 $f(A_1, A_2)=A_1A_2^*$. 以(10)～(13)代入(9)式. 由于 $A_j(j=1,2)$是随时间缓慢变化的函数，因此方程中随时、空迅变的相位因子应相消. 这样我们最后得到无吸收情况下二波耦合的方程：

$$\left.\begin{aligned}\frac{dA_1}{dz}&=-\frac{1}{2I_0}\Gamma A_2A_2^*A_1,\\ \frac{dA_2}{dz}&=\frac{1}{2I_0}\Gamma^*A_1A_1^*A_2,\end{aligned}\right\} \tag{14}$$

式中复耦合系数

$$\Gamma=i\frac{2\pi n'}{\lambda\cos\theta}e^{-i\varphi'}=\gamma+i2\beta, \tag{15}$$

参数 γ 和 β 分别为

$$\gamma=\frac{2\pi n'}{\lambda\cos\theta}\sin\varphi', \quad \beta=\frac{\pi n'}{\lambda\cos\theta}\cos\varphi'. \tag{16}$$

当我们引入线性吸收系数 α 后，(14)式可改写为

$$
\left.\begin{aligned}
\frac{\mathrm{d}A_1}{\mathrm{d}z} &= -\frac{1}{2I_0}\Gamma A_2 A_2^* A_1 - \frac{\alpha}{2}A_1, \\
\frac{\mathrm{d}A_2}{\mathrm{d}z} &= \frac{1}{2I_0}\Gamma^* A_1 A_1^* A_2 - \frac{\alpha}{2}A_2.
\end{aligned}\right\} \tag{17}
$$

(17)式对于光强 I_1 和 I_2 的解(参见附录 A9.1)为

$$
\left.\begin{aligned}
I_1(z) &= \frac{1+m^{-1}}{1+m^{-1}\mathrm{e}^{\gamma z}}\,\mathrm{e}^{-\alpha z}, \\
I_2(z) &= \frac{1+m}{1+m\mathrm{e}^{-\gamma z}}\,\mathrm{e}^{-\alpha z},
\end{aligned}\right\} \tag{18}
$$

式中

$$
m = I_1(0)/I_2(0), \tag{19}
$$

为输入端两束光强之比. 由(18)式可以看出，两个光束之间的能量传递情况与耦合系数的实部 γ(有时称 γ 为耦合系数)有很大关系. 由(16)式，当 $\varphi'=\pi/2$，即相位光栅与光强周期分布的相位差为 $\pi/2$ 时，γ 最大. 参数 γ 的符号取决于晶体光轴 c 的方向. 在 $\gamma>0$的那些方向上，随着光波的传播，光束 E_1 的能量逐渐传递给 E_2.

设 E_2 为信号光，E_1 为泵浦光，则增益系数定义为

$$
g = \frac{I_2(L)}{I_2(0)} = \frac{1+m}{1+m\mathrm{e}^{-\gamma L}}\,\mathrm{e}^{-\alpha L}, \tag{20}
$$

式中 L 是晶体的厚度. 在小信号近似下，$m\gg1$，则增益系数近似表为

$$
g \approx \mathrm{e}^{(\gamma-\alpha)L}, \tag{21}
$$

在 $\gamma>\alpha$ 的那些方向上，增益是正的，光束 2 得到放大. 根据朱镛等的报道[11]，对于掺 Ce 的钛酸钡 Ce:$BaTiO_3$，γ 在很大的光波传播方向范围内达到 30 cm^{-1}以上($\lambda=514.5$ nm)，而 α 约为3 cm^{-1}，即 γ 比 α 高出一个量级，因此在这些方向上光束 2 获得高倍放大. 在极端的情况下(γL 足够大)，输出端只有光束 2 而没有光束 1.

图 9.9 给出 $I_1(\gamma z)$(实线)和 $I_2(\gamma z)$(虚线)作为 γz 的函数，

设 $\alpha=0$，$m=100$，即小信号、无吸收近似. 随着光波在晶体中的传播及它们之间的能量耦合，I_2 逐渐增大，I_1 则逐渐变弱，能量从光束 1 逐渐传递到光束 2 中去.

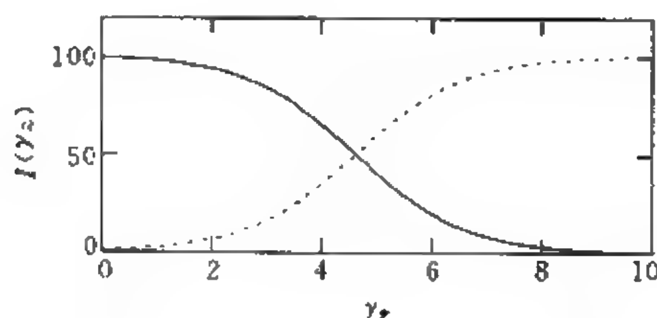


图 9.9　二波混频中的光能量的传递

这样一来，我们就看到了二波混频的这样一幅物理图象：两个光波在光折变晶体中相干叠加，形成周期性的光强分布. 这一光强分布又通过光折变效应在晶体中产生折射率的周期变化——相位光栅. 两个光波显然都自动满足 Bragg 条件，因而在它们自己产生的光栅上发生衍射：光束 1 的衍射光与光束 2 方向完全相同，光束 2 的衍射光与光束 1 方向完全相同，称为自衍射(self-diffraction)效应.

我们还可以看到相位光栅与光强光栅的相位差 φ' 的重要作用. 由(16)式，在 $\varphi'=0$ 时 $\gamma=0$，光束间是没有耦合的；$\varphi\neq0$ 时 $\gamma\neq0$，则能量在光束间传递，传递的方向取决于 γ 的符号. 在 $|\gamma L|$ 较大的情况下，一个光束在传播过程中，它的能量越来越多地转移到另一光束中去. 然而，与此同时，随着该光束强度的减弱，又影响了相干叠加光强条纹的调制度，从而影响相位光栅的调制度 $\Delta n/n$，这两个相互影响、相互制约的过程最后达到平衡. 在极端情况下泵浦光被信号光耗尽.

9.3.2　四波混频

让我们来考虑图 9.10 的情况：一对光束 E_1，E_2 反向射入晶

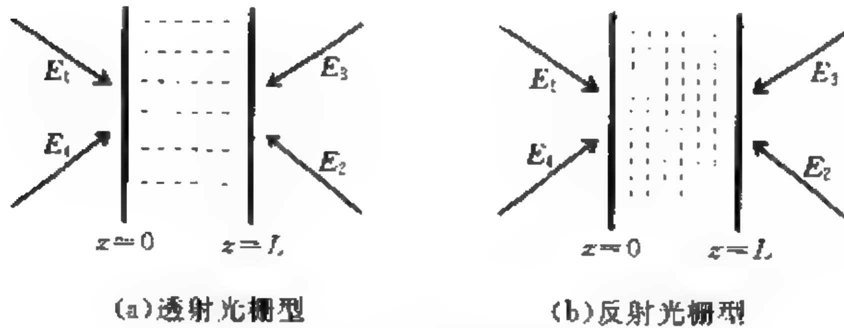
体，称泵浦光．第三个光波(信号光)E_4射入晶体后，三个光波共同作用的结果产生了第四个光波 E_3，它是信号光的相位共轭光．图 9.10(a)表示晶体中存在一个透射型光栅，四个波通过该光栅相互作用；图 9.10(b)则表示晶体中存在一个反射型光栅．我们仅限于考虑透射光栅的情况．设四个光束的波矢量均位于 xz 平面内，且有

$$\boldsymbol{k}_1=-\boldsymbol{k}_2, \tag{22}$$

$$\boldsymbol{k}_3=-\boldsymbol{k}_4, \tag{23}$$

$\boldsymbol{k}_1$ 和 $\boldsymbol{k}_4$ 关于 z 轴对称分布，在晶体中与 z 轴夹角均为 θ. 有

$$k_{1z}=k_{4z}\approx n_0k_0\cos\theta,\qquad k_{2z}=k_{3z}\approx -n_0k_0\cos\theta, \tag{24}$$



(a)透射光栅型　　(b)反射光栅型

图 9.10　四波混频

再设晶体中占主导地位的透射光栅是 E_1,E_4 及 E_2,E_3 形成的，可表为

$$n=n_0+\frac{n'}{2}e^{i\varphi}\frac{A_1^*A_4+A_2A_3^*}{I_0}e^{i\boldsymbol{K}\cdot\boldsymbol{r}}+\text{c.c.}, \tag{25}$$

式中 $\boldsymbol{K}$ 为光栅的波矢量，满足条件

$$\boldsymbol{K}=\boldsymbol{k}_4-\boldsymbol{k}_1=\boldsymbol{k}_2-\boldsymbol{k}_3, \tag{26}$$

$$f(A_1,\cdots,A_4)=A_1^*A_4+A_2A_3^*. \tag{27}$$

实现"单一光栅"的方案很多，例如可以控制相干长度，也可以选择入射光的偏振态等．将以上各式代入(9)式，得到四波混频情况下的耦合波方程

$$
\left.\begin{aligned}
\frac{\mathrm{d}A_1}{\mathrm{d}z} &= \frac{\Gamma}{2I_0}(A_1A_4^* + A_2^*A_3)A_4 - \frac{\alpha}{2}A_1, \\
\frac{\mathrm{d}A_2^*}{\mathrm{d}z} &= \frac{\Gamma}{2I_0}(A_1A_4^* + A_2^*A_3)A_3^* + \frac{\alpha}{2}A_2^*, \\
\frac{\mathrm{d}A_3}{\mathrm{d}z} &= -\frac{\Gamma}{2I_0}(A_1A_4^* + A_2^*A_3)A_2 + \frac{\alpha}{2}A_3, \\
\frac{\mathrm{d}A_4^*}{\mathrm{d}z} &= -\frac{\Gamma}{2I_0}(A_1A_4^* + A_2^*A_3)A_1^* - \frac{\alpha}{2}A_4^*,
\end{aligned}\right\} \tag{28}
$$

式中 α 为线性吸收系数，Γ 为复耦合系数，见(15)式.

我们进一步设

$$
I_1, I_2 \gg I_3, I_4, \tag{29}
$$

该条件称未耗尽近似，即泵浦光足够强，信号光和相位共轭光为微扰. 在四波混频过程中，信号光和相位共轭光未能显著地消耗泵浦光的能量. 略去 A_3, A_4 的二级小量，(28)式的解为(参见附录 A9.2)

$$
\left.\begin{aligned}
A_1(z) &= A_1(0)\mathrm{e}^{-\alpha z/2}, \\
A_2(z) &= A_1(L)\mathrm{e}^{\alpha(z-L)/2}, \\
A_3(z) &= \frac{A_2(L)A_4^*(0)\mathrm{e}^{\alpha(z-L)/2-\Gamma z/2}}{A_1^*(0)} \cdot \frac{J(z)}{J(0)+1}, \\
A_4^*(z) &= A_4^*(0)\mathrm{e}^{-\alpha z/2-\Gamma z/2}\frac{J(z)+\mathrm{e}^{\Gamma z/2}}{J(0)+1},
\end{aligned}\right\} \tag{30}
$$

其中

$$
J(z) = \int_{\Gamma z/2}^{\Gamma L/2} \frac{\exp\zeta \mathrm{d}\zeta}{1 + r\mathrm{e}^{-\alpha L}\exp(4\alpha\zeta/L)}, \tag{31}
$$

r 为泵浦光强度比，

$$
r = \frac{A_2(L)A_2^*(L)}{A_1(0)A_1^*(0)} = \frac{I_2(L)}{I_1(0)}, \tag{32}
$$

相位共轭反射率

$$
\rho = \frac{A_3(0)}{A_4^*(0)} = \frac{A_2(L)J(0)}{J(0)+1}\ \mathrm{e}^{-\alpha L/2}. \tag{33}
$$

如果吸收系数 $\alpha\approx 0$，那么 $J(z)$ 可以积出，从而有

$$
\left.\begin{aligned}
&A_1(z)=A_1(0),\\
&A_2(z)=A_2(L),\\
&A_3(z)=\frac{A_2(L)A_4^*(0)}{A_1^*(0)}\cdot\frac{e^{-\Gamma(z-L)/2}-1}{e^{\Gamma L/2}+r},\\
&A_4^*(z)=A_4^*(0)\frac{e^{-\Gamma(z-L)/2}+r}{e^{\Gamma L/2}+r}.
\end{aligned}\right\}\tag{34}
$$

无吸收情形下的相位共轭反射率

$$
\begin{aligned}
\rho&=\frac{A_3(0)}{A_4^*(0)}=\frac{A_2(L)}{A_1^*(0)\sqrt{r}}\frac{e^{\Gamma L/4}-e^{-\Gamma L/4}}{(e^{\Gamma L/4-(\ln r)/2}+e^{-\Gamma L/4+(\ln r)/2})}\\
&=\left[\frac{A_1(0)A_2(L)}{A_1^*(0)A_2^*(L)}\right]^{1/2}\frac{\sinh(\Gamma L/4)}{\cos[\Gamma L/4-(\ln r)/2]}.
\end{aligned}\tag{35}
$$

在文献[1]中用 $\gamma=-\Gamma/2$ 表示耦合系数(注意这里的 γ 与(15)式的定义不同)．以后，本章将沿用这一习惯定义，相位共轭反射率则表为

$$
\rho=-\left[\frac{A_1(0)A_2(L)}{A_1^*(0)A_2^*(L)}\right]^{1/2}\frac{\sinh(\gamma L/2)}{\cosh[\gamma L/2+(\ln r)/2]}.\tag{36}
$$

而相位共轭光强反射率则为

$$
R=\frac{A_1(0)A_2(L)}{A_1^*(0)A_2^*(L)}\left\{\frac{\sinh(\gamma L/2)}{\cosh[\gamma L/2+(\ln r)/2]}\right\}^2,\tag{37}
$$

γL 大体上相当于耦合系数和耦合长度的积，可称为耦合强度．图 9.11 给出在相同的 $\gamma L(|\gamma L|=3.627)$、不同的相位差 φ' 情况下，R 作为泵浦比 r 的函数．

我们注意到在 $\varphi'=0$(相位光栅与光强光栅同相)时，$r=1$(泵浦光等强)，得到最高的相位共轭光强反射率 R_{max}；$\varphi'\neq 0$ 时，R_{max} 对应的 $r\neq 1$. 其次，在常用的扩散极限($\varphi'=\pi/2$)下并未出现振荡．

φ' 从零逐渐增加，而相位光栅相对于光强光栅逐渐滞后，在 $\varphi'=\pi/6$，$r=6.13$ 时出现自激振荡($R\sim\infty$)，在有限的探测信号光强下获得极强的相位共轭反射光．当然，由于吸收的存在，R

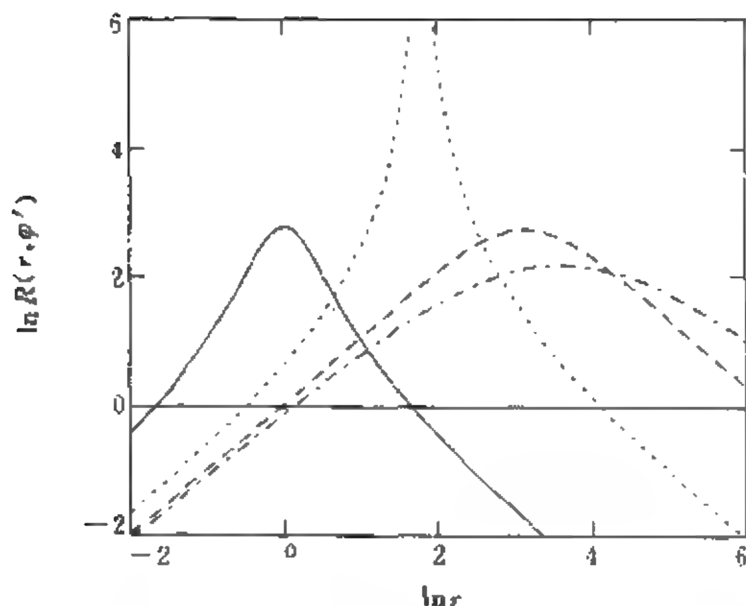


图 9.11 $|\gamma_L|=3.627$ 时的相位共轭反射率

纵坐标为 $\ln R(r,\varphi')$，横坐标为 $\ln r$. 自左至右，$\varphi'=0,\ \pi/6,\ \pi/3,\ \pi/2$

实际上不可能无限大. 图 9.12 为探测信号光 A_4 及相位共轭光 A_3 在晶体中传播过程中光强的变化，$\varphi'=\pi/6$，$r=5$. 我们注意到探测信号光强从入射面($z=0$)起逐渐增大，而相位共轭光在 $z=L$ 处为 0 并随着 z 的减小而加强，在入射面处相位共轭光强反射率已达 40 左右.

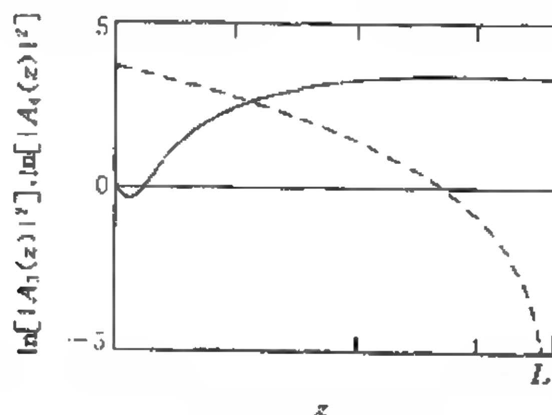


图 9.12 探测信号 A_4(实线)和相位共轭信号 A_3(虚线)在晶体中的光强分布

耦合波模型只是一级近似下光波传播和相互作用的简化模

型. 并且在以上的讨论中，都假定晶体是均匀一致的. 实际的晶体中还存在缺陷，此时光折变现象更加复杂.

当波矢量条件(22)、(23)严格满足，亦即 E_2 和 E_1 严格反向传播时，所谓"动量守衡条件"(26)式就能满足，这样一来，在入射面，E_3 和 E_4 的振幅互为复共轭，E_3 就是 E_4 的相位共轭波. 可以说，动量守衡是产生相位共轭的必要条件. 一般情形下，E_4 可能荷载空间信息，它不一定是平面波. 此时我们总可以把它按角谱(傅里叶谱)展开(参见第一章有关角谱的分析)，若每一角谱分量都满足条件(33)，E_3 仍然精密复制出 E_4(E_3 为 E_4 的相位共轭波). 这相当于要求导出(33)式的过程对于所有的傅里叶分量都是线性的.

9.4 自泵浦相位共轭效应

9.4.1 全内反射自泵浦相位共轭(total internal-reflection(TIR) self-pumped phase conjugation)

在节 9.3 中我们介绍了用四波混频在光折变晶体中产生相位共轭波的效应. 在这一装置中，用一对外来的泵浦光 E_1,E_2 照射晶体，且 $\boldsymbol{k}_1=-\boldsymbol{k}_2$，即泵浦光是严格反向传播的，否则就不能满足 Bragg 衍射的条件. 泵浦光与探测信号光 E_4 通过光折变效应互相耦合，最终产生 E_4 的相位共轭光.

外泵浦光并不一定必要，采用特殊的物理机构，信号光波可以自行产生相位共轭光，称自泵浦相位共轭效应(self-pumped phase conjugation，简写为 SPPC).

节 9.3.2 介绍了光折变晶体中的二波混频效应，我们指出，在二波混频中，两个光束在光折变介质中借助于它们自己产生的相位光栅而互相耦合，即自衍射效应. 在小信号近似下，光束的增益系数

$$g\approx e^{(\gamma-\alpha)L}, \tag{1}$$

在 $\gamma \gg \alpha$ 的那些方向上，

$$g \approx e^{\gamma L}, \tag{2}$$

γ 为耦合系数，L 是晶体厚度，即两个光束的作用长度，因而 γL 称为耦合强度. γ 依赖于光波在晶体中的传播方向，由节 9.3(16) 式可知，γ 是 θ 和 φ' 的函数. 在许多光折变晶体中，沿着某些方向 γ 相当大. 例如，当 $\gamma L = 20$ 时，g 竟高达 10^9.

当一束光射入晶体时，在晶体表面或内部的缺陷处可能发生光的散射. 散射光一般具有很宽的角谱，即光波是向各个方向散射的，而那些沿着高增益方向的散射光，将与原入射光波发生二波混频，从而产生高倍放大，这一效应称"扇出效应"(fanning).

图 9.13 所示为 Feinbeig 在 1982 年发现的著名的自泵浦相位共轭效应[9]. 所用的晶体为 $BaTiO_3$，光束 1 从左上角射入晶体，光轴从上向下，光波为非常光. 设在晶体内 A 处，由于缺陷发生散射，散射光 2′ 在晶体内被它的两个相邻表面反射两次后反向传播，成为光束 3，并与入射光相交于 B 处，此时相交成 90° 的两个表面的作用相当于光束从直角棱镜中的两次全反射.

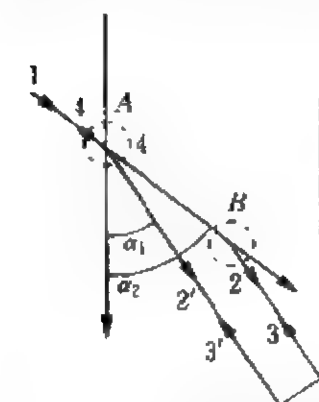


图 9.13　用晶体的一条棱作为内反射器的 SPPC

B 处的散射的光束 2 则沿着相反方向通过上述回路. 光束 1 和 2′ 经过二波混频形成透射光栅，而反向传播的散射光 3′ 则被这一透射光栅衍射，衍射光 4 与原光束 1 的方向相反，亦即光束 3′ 读出由光束 1 和 2′ 形成的光栅，构成光束 1 的相位共轭光. 同样，入射光束 1 与散射光束 2 经二波混频形成相位光栅，而光束 3 则"读"这一光栅，形成相位共轭光 4. 事实上，在 A、B 两个区域内，各光束是同时存在的，由于两个相交平面的反射，形成了与散射波反向传播的光束，通过四波混频效应，产生相位共轭光.

显然，从入射光扇出的各个方向的散射光并不都能构成上面

的振荡回路，只有沿着增益最大的方向或接近此方向的那一小部分散射光才能形成振荡，其他方向的散射光都因增益不够而不能自举.

节 9.3(16)式表明耦合常数与 n' 成正比. 由节 9.2 可知，n' 与有效线性电光系数 γ_{eff} 成正比，因此耦合效应的强弱也就取决于 γ_{eff} 的大小. 在 $BaTiO_3$ 中，e 光的 γ_{eff}(参见节 9.2(30)式)

$$\gamma_{\text{eff}}=\frac{1}{n_e n_o^3}\Big[n_e^4\gamma_{33}\sin\alpha_1\sin\alpha_2+2n_e^2n_o^2\gamma_{42}\cos^2\frac{\alpha_1-\alpha_2}{2}$$

$$+n_o^4\gamma_{13}\cos\alpha_1\cos\alpha_2\Big]\sin\frac{\alpha_1-\alpha_2}{2},\tag{3}$$

其中 α_1 和 α_2 见图 9.13.

$BaTiO_3$ 晶体和掺铈 $BaTiO_3$ 晶体($Ce:BaTiO_3$)的耦合系数 γ 与 α_1,α_2 的函数关系如图 9.14 所示，从左到右各曲线对应的 α_1 分别为 15°，30°，45°，60°，75°，晶体中载流子浓度 $N=3\times10^{16}\text{m}^{-3}$，入射光为 e 光，耦合强度 $\gamma L\geqslant 2.34$. 可以看出，α_1 从 15°到 75°的范围内都有明显的效应，且在 α_1 和 α_2 的差约为 5°时达到极大值. 在实验中，当入射角改变时，回路方向自动改变，跟踪入射光，使散射光永远处于增益为极大的方向. 当然该回路必须将晶体的两面角的交棱包含在内. 这一自动跟随、自动调准的性能充分证明了上面的分析是正确的.

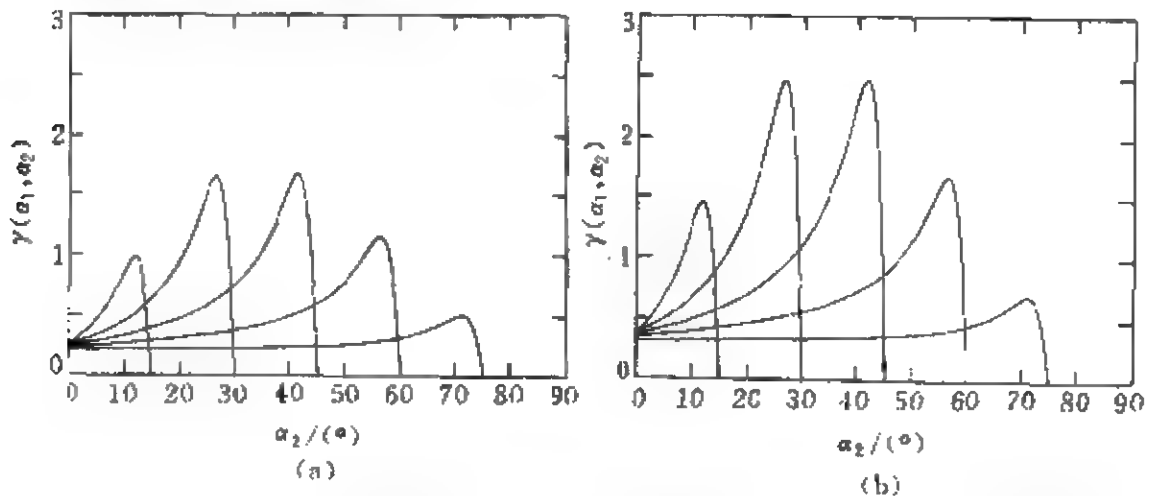


图 9.14 (a) $BaTiO_3$ 中耦合系数 γ 作为 α_1,α_2 的函数 (b) $Ce:BaTiO_3$ 的相应曲线

当入射光是高斯型激光时，图 9.13 中的回路比较简单；当入射光被一个物体调制时，例如在入射光中插入一个透明的输入图形，回路就出现分叉，但各分叉的走向大体一致．这是因为入射光受到图形的调制时，可以把它按角谱展开，每一角谱分量都是平面波．各平面波分别形成各自的相位光栅，这些相位光栅的方向略有差别，因此各自的振荡回路也略有差别．

Feinberg 首先用实验证明了上述结论[9]，参见图 9.15．一束准直的激光 A_r^+ 激光(λ=514.5 nm，TEM_{00}模)通过一张猫的照片，再用 f=250 mm 的透镜聚焦到 $BaTiO_3$ 晶体中，功率约 5 mW．用一个分光镜引出相位共轭波，照到一个屏上．光束的入射角(光束与晶体表面夹角)在 10°到 80°的大范围内任选．把照片插入光束后，相位共轭像延迟数秒后出现．当照片缓慢地平移时(速度为几 mm/s)，相位共轭像也同步连续移动．当照片快速位移时，原来的共轭像逐渐消失，并出现在另一位置上，这一过程约几秒钟．这一时间与 $BaTiO_3$ 晶体中相位光栅的建立过程的弛豫时间大体相当．

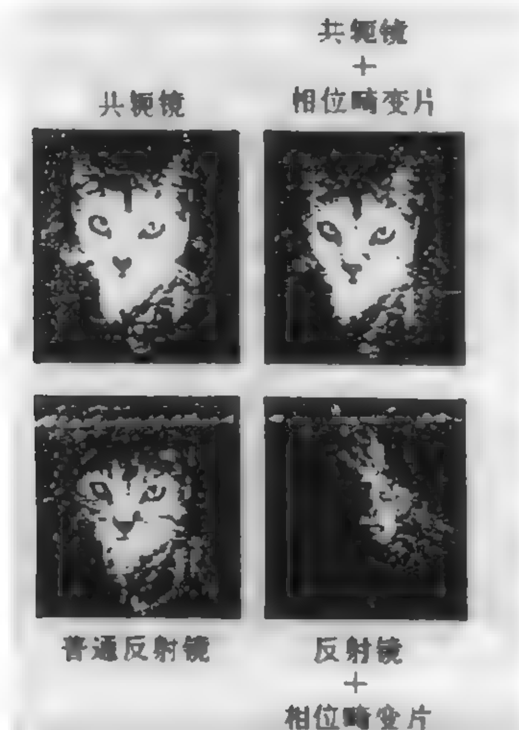


图 9.15 “猫”相位共轭[9]

以下的实验是很有意义的：当在物体和晶体之间放置一个随机相位畸变板时，相位共轭像仍然十分清晰，这是因为相位共轭光波具有相位恢复的功能，这也是 SPPC 的反射波即为入射信号波之相位共轭波的最有力的证明.

自泵浦相位共轭具有装置紧凑、反射效率高、入射角范围大、振荡回路自动跟随入射光等特点，自泵浦生成的相位共轭像清晰度高，特别是由于参与混频的光波都源于单一入射光，因此，它对于防振的要求不高.左上图为猫的照片经 SPPC 形成的相位共轭像;左下图为普通反射镜形成的反射像. 分别在晶体前和普通反射镜前加一片相位畸变片，由于相位恢复效应，相位共轭像仍很清晰，见右上图;但普通反射镜的反射像已完全失真，见右下图.

由于该实验装置中输入图像是猫，因而又称为"猫"共轭(cat conjugation)，在晶体中通过全内反射形成振荡回路则是它的显著特征.

9.4.2 后向散射自泵浦相位共轭

在适当的条件下，一束射入晶体的激光，在它传播的路径上，同样由于缺陷而引起的后向散射光可能与入射光形成二波混频，产生反射形光栅，见图 9.16(a). 该光栅的波矢量 $K=2k$，又称 $2k$ 光栅. 该效应称后向散射自泵浦相位共轭，又称受激光折变后向散射(stimulated photorefractive backscattering，SPB)模型. 这一模型的机理将在下一节中进一步探讨. 按目前的认识，产生相位共轭的效应仍然是四波混频，但在晶体中反射光栅与透射光栅同时存在. 入射光既是泵浦光，又是读出光. 入射光波自己读这一相位光栅，产生了相位共轭波，就像光波从普通的反光镜中反射一样，只是反射光波是从原路折返的.

后向散射 SPPC 与全内反射 SPPC 形成的机理不同，它不需要借助于扇出光在晶体内表面的反射，因此只需要把入射面抛光

就足以引起相位共轭．通常入射面取为(100)面，它更像是普通的反光镜，因此该器件又称自泵浦相位共轭镜(self-pumped phase conjugator mirror，SPPCM)．

由于耦合系数的大小主要取决于晶体的有效电光系数 γ_{eff}，所以人们一直在寻求具有大电光系数的晶体．朱镛、吴星等[10,11]报道了由中国科学院物理研究所研制的掺铈 $BaTiO_3$(即 Ce：$BaTiO_3$)，具有比常规的 $BaTiO_3$ 更高的 γ_{42}(见表 9.1)，所以在相当大的入射角范围内获得了很高的自泵浦相位共轭反射率．

表 9.2 给出 0°，45°切割的 $BaTiO_3$ 及 Ce：$BaTiO_3$ 的 SPPC 反射率 ρ 及最高反射率 ρ_{max}，其中 45°切割 Ce：$BaTiO_3$ 在 514.5 nm 下的 SPPC 反射率达 83.5%，是目前国内外的最高水平．

表 9.1 Ce：BaTiOS₃ 的线性电光系数(λ=633 nm)[11]

Ce×10⁻⁶	γ_{13}/(pm/V)	γ_{33}/(pm/V)	γ_{42}/(pm/V)
0	8±2	105±10	1300±100
15			1880±150
26			1845±150
33	11.8±1.0	113±8.5	1945±220

表 9.2 $BaTiO_3$ 及 Ce：$BaTiO_3$ 的 SPPC 反射率(λ=514.5nm)[11]

样品	建立时间/s	入射角范围/(°)	ρ/(%)	ρ_{max}/(%)
0° $BaTiO_3$	14	+60～+70	>40	50
45° $BaTiO_3$	7	+40～+75	>60	75
0° Ce：$BaTiO_3$	8	+35～+70	>60	70
45° Ce：$BaTiO_3$	3	−60～+70	>60	83.5

图 9.16(b)为后向散射自泵浦相位共轭镜产生的相位共轭图像，可以看出图像的保真度非常高．图(c)表示在晶体前插入相位畸变板后的相位共轭图像，仍有较高的分辨率．图(d)是用一块普通反光镜代替晶体，并在反光镜前插入相位畸变板(像差片)后

的反射像，我们看到物体的像已荡然无存．张洪钧等对自泵浦相位共轭保真度进行了深入的研究，有兴趣的读者请参阅他们的论文[48,49]．

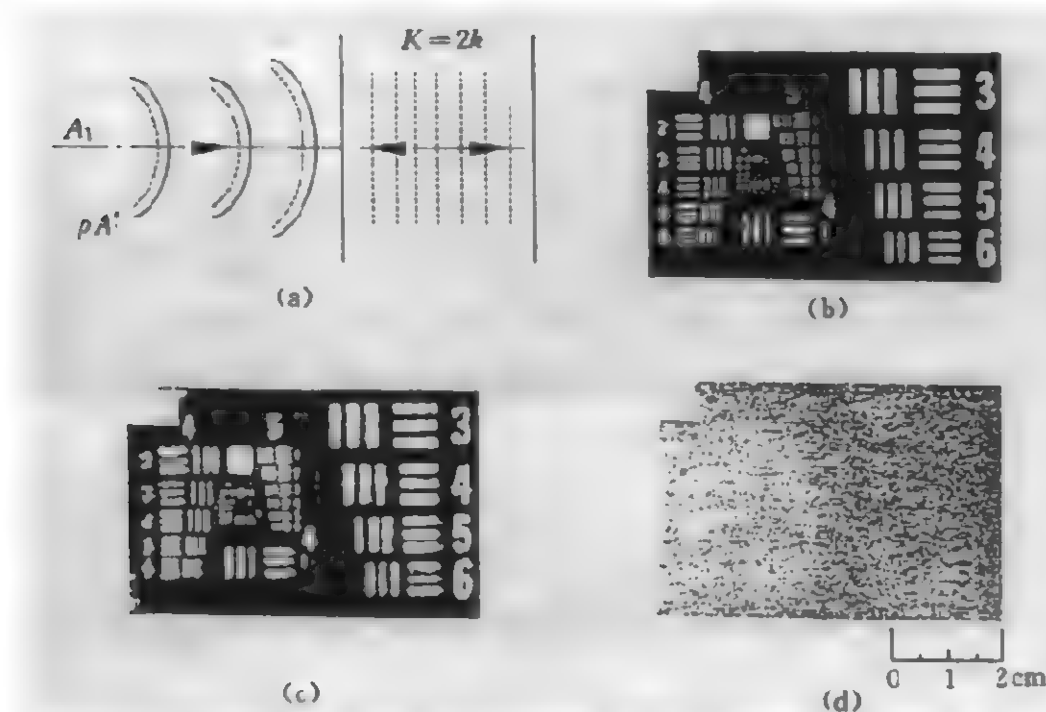


图 9.16 受激光折变后向散射(SPB)相位共轭图像及其保真度[10]

(a) 受激光折变后向散射(SPB)示意图 (b) 没有像差片 (c) 带有像差片

(d) 用一块普通反光镜代替晶体，并在反光镜前插入像差片后的反射像

9.4.3 对于 SPPC 机理的进一步研究

自从 Feinberg 在 1982 年首次观察到自泵浦相位共轭效应以来，至少已有三个模型，用来解释 SPPC．第一个是 McDonald 和 Feinberg[9,28] 提出来的全内反射模型(total interal reflection，TIR)，用来解释"猫"共轭的成因，我们已在节 9.4.1 中作了介绍．在这一模型中，在晶体内的两个区域中发生四波混频效应．而它们彼此间又通过光波在晶体内相交的两个表面的内反射耦合，形成回路，产生自举，这也是该模型的特征和必要条件．第二

个模型是受激光折变后向散射模型(stimulated photorefractive backscattering, SPB),我们在节9.4.2中作了介绍. 在该模型中,入射光在它的传播路径上由于缺陷等原因产生的后向散射光与前向入射光形成二波混频,产生2k光栅. 入射光自己读这一光栅产生相位共轭波,它的特点在于,入射波既是泵浦光又是读出光,SPPC的产生不必借助两次内反射,晶体只需入射面抛光,即可产生SPPC效应. 第三个模型是在掺铁KTN和掺铈$BaTiO_3$晶体的SPPC效应的研究中提出来的,称为光折变后向散射-四波混频模型(SPB-FWM),在这一模型中,入射光的扇出光和扇出光激发出的后向散射光构成一对反向传播的光束,它们和入射光一道,在晶体内的一个(而不是两个)区域内通过四波混频效应产生相位共轭光. 这一模型不需要内反射.

叶佩弦等[29~34]深入研究了这三个模型,探讨了它们间的关系和互相转化的条件,提出了一个统一的模型[35]来解释表面上似乎并不相同的三种生成SPPC的机制,深刻地揭示了这一效应的本质.

首先,他们研究了在波长不相同,其余条件完全相同的情况下SPPC的行为,他们发现在波长较短时,SPPC是通过SPB-FWM机制产生的,而在较长的波长下,SPPC必须借助于TIR才能产生. 这是因为散射截面显然随波长的变短而迅速变大. 在短波情况下,散射效率很高,光折变后向散射波足够强,在一个区域内通过四波混频就产生了相位共轭波.

他们又保持其他条件不变,仅改变掺入杂质的浓度来观察SPPC,发现随着杂质浓度的下降,SPPC逐渐从SPB-FWM向TIR转化,亦即两种模型之间的转换波长随掺杂浓度的下降而发生蓝移现象. 这是因为随着掺杂浓度的下降,光折变的的总体效率下降,只有较短波长的光产生的散射光,才足以建立后向散射四波混频.

图9.17给出不同的掺杂浓度和波长对SPPC模型的影响示

意图，该图表明，在较短的波长或较高的杂质浓度下，一般比较容易通过后向散射-四波混频形成相位共轭，而在较长的波长或较低的杂质浓度下，散射截面较小，只能借助全内反射产生相位共轭光，即"猫"共轭. 在两个模型之间则是过渡区.

叶佩弦等认为，纯粹的后向散射模型(SPB)并不能产生相位共轭光，亦即四波混频(FWM)是产生相位共轭的必要条件，SPB也必须通过FWM来产生相位共轭光. 因此，对于SPPC而言，必须存在一个产生反向传播光束的根源或机制，不同的机制对应于不同的模型，SPB-FWM或TIR. 也就是说，不同的模型间的差别仅在于反向传播光束的产生方式，但本质是相同的，都是四波混频.

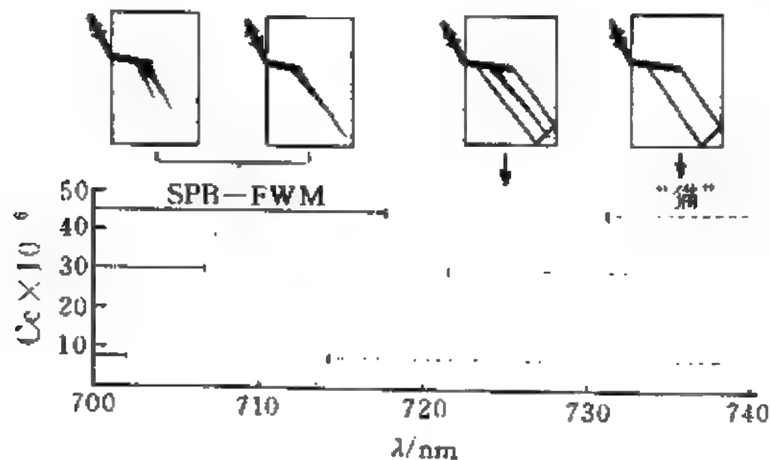


图 9.17　SPPC 机制转换波长范围与 Ce 在 $BaTiO_3$ 中掺杂浓度的关系[35]
"——"表示 SPB-FWM 机制，"------"表示 TIR 机制，中间为过渡区

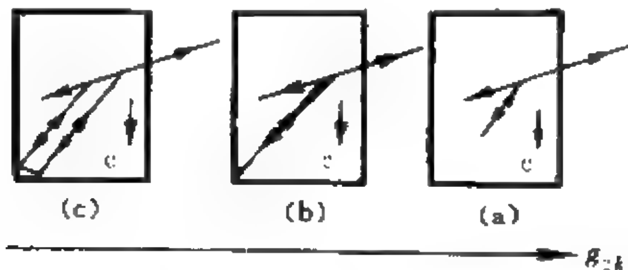


图 9.18　SPPC 机制转换与后向散射增益系数 g_{2k} 的关系[35]

如果散射截面足够大，扇出光的后向散射光足够强，则 SPPC 通过一个区域内的 SPB-FMW 效应就建立起来；否则，如后向散射光不够强，则必须借助于全内反射(TIR)，在两个区域内产生 FMW，通过一个回路互相耦合，产生自举，建立 SPPC.

如果用 g_{2k}来表征后向散射的增益系数，则 SPPC 的建立机制和 g_{2k}的关系如图 9.18 所示. 在 g_{2k}较大时(a)，扇出光在晶体中较短距离的传播过程就激发出足够强的后散射光，通过 SPB-FMW 建立 SPPC；随着 g_{2k}的减小(b)，扇出光必须通过较长距离的传播，才能激发出足够强的后向散射光，通过 FMW 建立 SPPC；在 g_{2k}继续减小的情况下(c)，单单依靠后向散射光并不能建立起能够自举，即能持续进行的 SPPC，反向传播光机制就从 SPB-FMW 转换成 TIR.

图 9.19 给出二波耦合增益系数 Γ_{2k}和波长的关系. 杂质的作用则是增加在晶体中有效电荷载流子的浓度，从而在一定的波长下使后向散射的增益变大，因此在掺杂浓度较高的光折变晶体中容易观察到后散射-四波混频(SPB-FWM)自泵浦相位共轭效应.

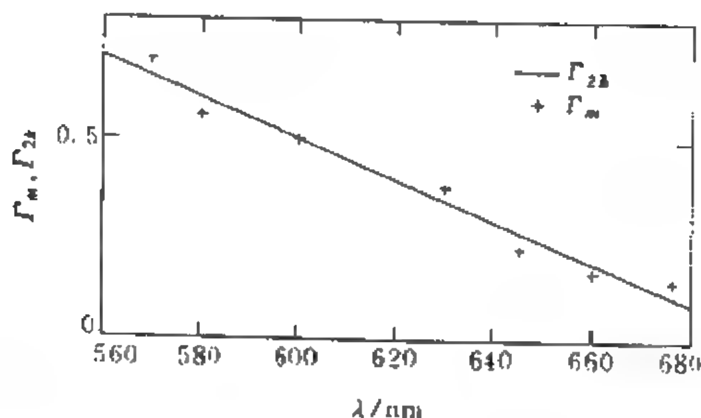


图 9.19　二波耦合增益系数 Γ_{2k}与波长的关系[35]

9.4.4　扇出效应

扇出效应(funning effect)，指的是入射到光折变晶体中的光

束，由于晶体中缺陷和杂质的散射，在传播过程中逐渐形成扇形的光强分布. 这是光折变晶体中光波传播的一个颇具特色的效应. 在自泵浦效应中，扇出效应是至关重要的. 然而在另一些应用中，扇出效应却引起了入射光的损耗，影响了二波混频的增益，应设法加以避免[39-42].

由于扇出效应对光波在光折变晶体中的行为有重要的影响，因此有关这一效应机理的研究也在进行. Feinberg 等[12]提出自散焦机构(self-defocusing)，而 Valley 等[43]则认为扇出效应是由于散射光通过与入射光的二波混频获得放大，并形成扇形分布. Gu 等[44]分析了光折变晶体中带电粒子的随机分布通过 Pockels 效应引起折射率光栅的随机噪声，散射截面则与折射率的起伏有关，散射光与入射光的耦合引起了扇出效应. 在上述各个模型中，仅考虑一个方向的散射光与入射光的相互作用.

张洪钧等[45]提出了一个改进模型，在这一模型中，考虑了各个方向的散射光与入射光的相互作用，他们假设散射光相对于入射光为小量，因而各散射光分量之间的耦合可以略去，仅考虑散射光和入射光的耦合，在这样的假设下建立了耦合波方程，计算所得的输出光强度的角度分布与 45°切割 $BaTiO_3$ 的实验结果符合得很好. 他们还讨论了扇出光波的瞬态特性[46].

在最近的一篇文献中，张洪钧等[47]深入研究了扇出光路的弯曲现象，他们首先考虑了输入光路的有限宽度，因此光场不仅是 z 的函数，还是 x 的函数. 他们还计入了扇出光各分量间的相互作用，是一个更加精确的二维模型.

他们预言，在足够长的传播路径中，已散布在一个相当大的角度范围内的扇出光将逐渐集中到一个"管道"中，即趋向于某一方向，角度分布趋近于 δ 函数，并最终沿光轴 c 传播，其他方向的散射光分量都将散逸. 这一效应的潜在应用，是制作一个激光导向器件(laser beam steering)：以任何方向射入晶体的激光束，只要晶体尺寸足够大，最后都将沿光轴传播.

在下一节互泵浦相位共轭效应中，我们将进一步看到扇出效应和扇出光弯曲的重要意义.

9.5 互泵浦相位共轭及双色泵浦光折变振荡

9.5.1 互泵浦相位共轭效应(MPPC)

继自泵浦相位共轭效应之后，Sternklar 等[15]、Sharp 等[13]、James 等[14]又先后发现了互泵浦相位共轭效应(mutually pumped phase conjugation，MPPC 或 double-phase conjugation，DPC)．典型的互泵浦相位共轭效应如图 9.20 所示．两个激光束 E_1,E_2 分

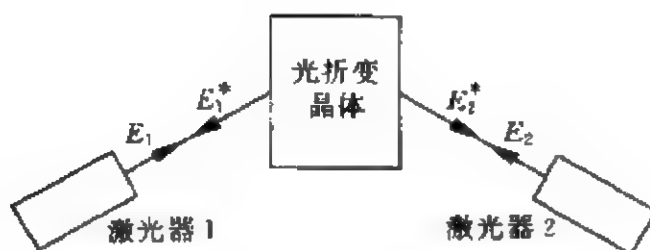


图 9.20 互泵浦相位共轭效应

别由两个不同的激光器辐射，从两侧射入光折变晶体，在恰当的几何配置下，同时产生它们的相位共轭波 E_1^* 和 E_2^*.

MPPC 效应具有以下显著的特点：

(1) 两个光束可以是来自不同激光器的激光，因而是不相干的，其波长一般相同，即 $\lambda_1=\lambda_2=\lambda$. 在 $\lambda_1\neq\lambda_2$ 时，反射光的波长与另一侧入射光的波长一致，而与同侧入射光的方向并不严格相反，有一个小的夹角，这一效应又称双色泵浦光折变振荡(double-color-pumped photorefractive oscillator，DCPO)．Sternklar 等[16]、Rabbani 等[18]详细讨论了该效应.

(2) 任何一个入射光本身并不产生相位共轭波，即不存在自泵浦相位共轭波.

(3) 任何一个入射光的相位共轭波都是另一个入射光泵浦的，这正是互泵浦相位共轭的主要特点，即 E_1^* 的能量来自 E_2(E_1^* 由 E_2 泵浦)，但 E_1^* 中包含 E_1 的空间信息. 换句话说，每一个相位共轭波的能量来自另一侧的入射光波，而空间信息来自同侧的入射光波.

(4) 两个相位共轭波同时产生，互不相干，亦即在稳态下，E_1^* 中完全没有 E_2 的信息，E_2^* 中也完全没有 E_1 的信息.

上述特点，使互泵浦相位共轭效应找到许多应用，特别是图像信息的转换，包括在两个光束之间的转换、相干态的转换等.

下面首先讨论这一效应的机理，主要是作者的观点，并参考了 Anderson 等[23]和 Sharp 等的论文[13].

9.5.2 MPPC 效应的机理

在节 9.4 中，我们介绍了光折变晶体的自泵浦相位共轭效应(SPPC). 产生这一效应的直接原因是入射光 E_1 波在晶体表面及内部缺陷上的散射，称为扇出效应，在扇出光波达到最大增益的那些方向上，入射光和散射光 E_{1f} 发生二波混频，产生相位光栅，入射波由它自己产生的相位光栅衍射，使能量转移到散射光中去. 形成散射光的放大. 这种情况下，只要存在与散射光反向传播的另一光波 E_{2f}，不论该光波是如何产生的，由于它满足 Bragg 条件，所以它读相位光栅并产生 E_1 的相位共轭波 E_1^*.

在全内反射型 SPPC(TIR-SPPC)中，E_{2f} 是散射光 E_{1f} 在晶体内两面角反射两次形成的. 如果 E_{2f} 来自晶体外部，就形成互泵浦相位共轭器 MPPC. 参见图 9.21(a)，设 E_1 在晶体内 A 点发生散射，并设 AB 方向大体上是散射光增益最大的方向，因此在该方向上的散射光 E_{1f} 与 E_1 通过光折变效应生成相位光栅 G_1，E_{1f} 借助于二波混频得到放大，或者说 E_1 泵浦晶体而产生 E_{1f}.

我们采用节 9.2.3 的全息光栅模型，E_1 和 E_{1f} 引起的折射率改变量可表为

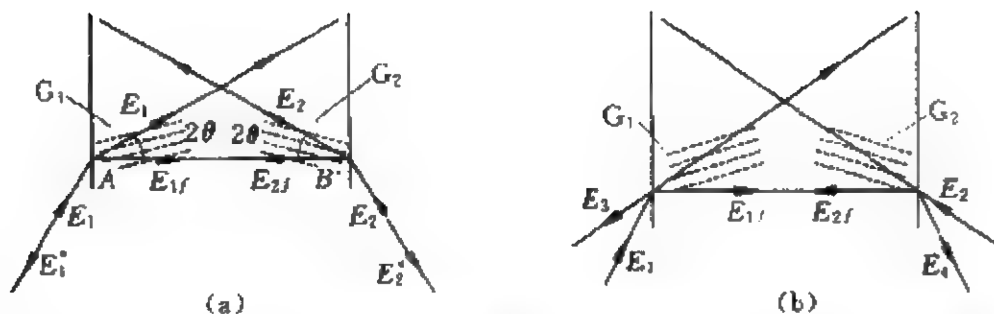


图 9.21[23]　(a) MPPC 的机理示意图　(b) DCPO($\lambda_1 \neq \lambda_2$)的机理示意图

$$\Delta n_1 = \frac{n'}{2} e^{i\varphi'} \frac{A_{1f} A_1^*}{I_1} e^{-i(k_1 - k_{1f}) \cdot r} + \text{c. c.}, \tag{1}$$

式中 A_1 和 A_{1f} 分别是 E_1 和 E_{1f} 的振幅，而 I_1 则是 E_1 和 E_{1f} 能量之和.

设晶体光轴是对称分布的，当晶体另一侧有一个对称入射的光波 E_2 时，E_2 同样会通过扇出效应而散射. 在 B 点散射的光波中，沿 BA 方向的散射波 E_{2f} 既获得放大，又满足 Bragg 条件，可记为

$$E_{2f} = A_{2f} e^{-i(k_{1f} \cdot r + \omega t)}, \tag{2}$$

E_{2f} 读光栅 G_1(被 G_1 衍射)而激发出一个新的波 E'，由节 9.2，该波可表为

$$E' \propto \Delta n_1 E_{2f} = C n' e^{i\varphi'} (A_{1f} A_{2f}) A_1^* e^{-i(k_1 \cdot r + \omega t)} + \text{c. c.}, \tag{3}$$

式中 C 为常数，E' 显然就是 E_1 的相位共轭波 E_1^*(差一个常数因子).

我们对上面的过程作进一步的讨论：

(1) 相位光栅的产生

我们把 E_j 和 E_{jf}($j=1,2$)经二波混频产生相位光栅的过程看成是全息记录的过程. E_j($j=1,2$)为物光，它们可能荷载输入图像的空间信息，一般具有复杂的波前. 扇出效应是一个复杂的散射过程，散射源的分布是随机的，从一个点散射的光波接近于球面波，因此扇出光 E_{jf}($j=1,2$)并不携带物光的信息，不妨把它看作

参考光. 根据上面所描述的过程，E_{1f}读光栅 G_2 而产生 E_2^*，E_{2f}读光栅 G_1 而产生 E_1^*. 既然 E_{1f}，E_{2f}都不携带它们所对应的物光的信息，所以 E_1^* 只有 E_1 的信息，E_2^* 只有 E_2 的信息，彼此间没有干扰.

(2) 相位共轭光的建立

在(3)式中，A_{1f}，A_{2f}分别是 E_{1f}，E_{2f}的振幅. 如果它们分别是一个点源散射的球面波，那么 A_{1f}，A_{2f}大体上是一个均匀的振幅分布，(3)式中 $Cn'e^{i\varphi'}(A_{1f}A_{2f})$可以近似认为是常数，因此 E' = 常数 $\times[A_1\exp(i\boldsymbol{k}_1\cdot\boldsymbol{r})]^*\exp(-i\omega t)$，它显然就是 E_1^*. E_{2f}读光栅 G_1 而产生 E_1^* 的过程，相当于全息再现中用反向传播的参考光来照射全息图. 在物光的相反方向得到实像，它正是物光的相位共轭像. 从四波混频的观点来解释，E_1，E_{1f}和 E_{2f}通过四波混频效应产生相位共轭波 E_1^*.

(3) 相位共轭光的信息源和能量源

E_{2f}读光栅 G_1 而产生 E_1^*(或者说 E_{2f}泵浦光栅 G_1 而产生 E_1^*)，G_1 是 E_1 中信息的载体，E_1 只是间接的信息源. 不论 E_1 是否存在、是否已发生变化，只要 G_1 没有变，E_{2f}在 G_1 上的衍射产生的仍然是原来的 E_1^*. 因此，E_1^* 的直接信息源是 G_1，而能量源是 E_{2f}. 上述结论对 E_2^* 同样适用.

(4) 正反馈和振荡的自举

再考虑一下图 9.21(a)的情况. 当 E_1 和 E_{1f}建立相位光栅 G_1 后，只要存在一个 E_{2f}，它的传播方向与 E_{1f}恰相反，则 E_{2f}对于 G_1 满足 Bragg 条件并发生衍射，产生 E_1 的相位共轭光 E_1^*. 由于 E_1^* 的传播方向与 E_1 相反，所以 E_{2f}和 E_1^* 通过光折变效应产生的相位光栅与 G_1 完全相同，亦即 G_1 得到加强. 其结果，更多的能量从 E_1 通过 G_1 的衍射转移到 E_{1f}中去. E_{1f}的强度增大，通过在 B 处 G_2 上的衍射使 E_2^* 强度加大，这样一来，G_2 同样得到加强，反过来又使更多的能量从 E_2 转移到 E_{2f}. 这是一个正反馈的过程，其

结果，使 AB 间建立振荡，获得很强的相位共轭输出 E_1^*，E_2^*. 其他方向的扇出光由于不满足 Bragg 条件，形不成正反馈，因吸收而变弱，最后在晶体中散逸.

当然，AB 方向上必须具有足够的增益，即 $\gamma-\alpha$ 足够大，才能形成振荡. 此外，在上述 E_{1f}，E_{2f}以及 E_2^*，E_1^* 增强的过程中，E_1，E_2 的能量被消耗而变弱，从而引起 E_1，E_{2f}相互作用产生的 G_1 以及 E_2，E_{2f}相互作用产生的 G_2 变弱. 这两个过程相互影响，相互制约，最后达到平衡. 所以说，互泵浦相位共轭的效率(相位共轭光强度与入射光强度的比，即互泵浦相位共轭镜反射率)R 取决于入射光角度、晶体光轴方向、E_{1f}和 E_{2f}所对应的 γ 及晶体的吸收系数 α 等.

(5) 能量转移、耦合的择优方向和桥共轭

扇出效应是一个复杂的多次散射过程，一般来讲，晶体内部及表面有足够稠密分布的散射源(杂质、缺陷)，一次散射的微弱散射光，通过与相对来讲很强的入射光的二波混频，从入射光吸收光能而获得放大，经过放大的一次散射光又可能发生再次散射. 最后，散射光的总能量变得足够大，形成的光栅也极为复杂. 在极端的情况下，入射光在晶体的传播过程中被逐渐耗尽，在输出端看不到入射光(或很微弱)，其能量在多次散射-衍射的过程中逐渐转移到散射光中去.

由于耦合系数 γ 是方位角 θ，φ 的函数，它和 α 的共同作用，制约了散射光的发展方向. 根据图 9.14，我们可以设想第一次散射光的主极大与入射光的夹角大约为 5°，第二次散射光的主极大与第一次散射光的主极大的夹角也是 5°，等等. 如晶体中散射源的分布充分稠密，则多次散射的主极大由折线变成凸曲线，大体上沿 γ 的极大方向，互相向对方弯曲. 一旦两条曲线互相搭接，正反馈立即建立，形成桥状的“增益回路”，又称桥共轭器(bridge conjugator)，上述过程在晶体中是自动完成的. 在 A，B 间可能有多个回路，但事实上桥一定沿着增益最大的回路建立，自动满足 Bragg

条件. 这一点颇有些像不均匀介质中的光路,该光路一定沿着光程最短的路径建立(称为费马原理). 如果我们用泛函来表征桥共轭问题,那么桥回路必然沿着使泛函取极值的回路建立. 也就是MPPC的建立过程具有自适应的特点,这一点与SPPC相似. 图9.21只画出一次散射过程,是桥共轭器的初级近似模型,但它的结论可以推广到更一般的情况.

Sharp等[13]曾做过实验,发现 E_1 和 E_2 的分布大体对称时效率最高;当 E_1 和 E_2 的分布偏离对称时,相位共轭信号迅速变小. 其次,他们又用一束与图9.22中 AB 方向平行的片状o光作为相位光栅的擦除光从上向下移动,发现当擦除光在 E_1 和 E_2 在晶体中的交点 O 以下、入射点 A,B 以上,擦除效果最为明显,可见 E_{1f} 和 E_{2f}的确在晶体中弯曲成桥状. 此外,相位共轭反射率还和入射光的入射角有关. 这一点已在SPPC中解释过了,参见图9.14及相应的说明. 当入射角变化时,上述桥状增益回路也自动跟随入射光路而变化,但相位共轭反射率会随之变化.

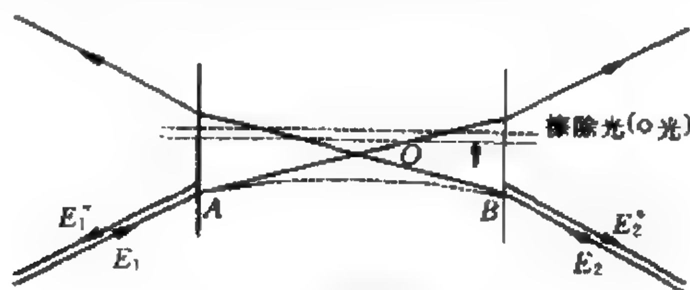


图 9.22 擦除效应[13]

(6) 相干性要求

正由于 E_{2f}仅仅是 E_1^* 的能量源,所以 E_2 不必与 E_1 相干. 亦即用另一相同波长激光器辐射的激光 E_2,就可以精确地复制出 E_1^*,E_1^* 的复制仅依赖于 E_{2f}读光栅 G_1 的过程. 这样一来,E_1^* 和 E_2^* 也是互不相干的. 然而,在晶体中形成的复杂的相位光栅,却对四个光波都起作用,它们不仅在自己形成的光栅上衍射,也在对方形成的光栅上衍射. 所以我们说,这四束光通过共有的光栅相

互作用，最后产生相位共轭效应.

(7) 波长偏调和双色泵浦光折变振荡器(DCPO)

当 E_1 和 E_2 的波长 $\lambda_1 \neq \lambda_2$ 时，E_{1t}读 G_2 得到 E_4，它与 E_2 的方向有一个小的夹角；E_{2t}读 G_1 得到 E_3，与 E_1 有一个小的夹角，参见图 9.21(b)，这一夹角显然是由于对 Bragg 条件的偏离形成的. 此时 E_3 并不是 E_1 严格的相位共轭光，E_4 也不是 E_2 严格的相位共轭光，这一情况类似于用 Ar^+ 激光去再现用 He-Ne 光记录的全息图，称为双色泵浦光折变振荡器(double-color-pumped photorefractive oscillator，DCPO). 如果把 $\lambda_1 = \lambda_2$ 的情况称为简并的四波混频，则当 $\lambda_1 \neq \lambda_2$ 时简并解除.

9.5.3 互泵浦相位共轭效应的稳态特性和瞬态特性

在晶体中建立相位光栅及稳定的桥型回路需要较长的时间，时间常数 τ 取决于晶体的种类和泵浦条件(光强、入射角、泵浦光的相干性等)，与入射光的截面尺寸也有关系.

上面讲相位共轭输出光 E_1^* 和 E_2^* 没有干扰，意思是说，尽管 E_1^* 是由 E_2 泵浦的，但它只从 E_2(通过 E_{2t})吸收能量，但 E_1^* 中并不携带 E_2 的信息，E_2^* 中也不带 E_1 的信息. 然而，这只是稳态的情况，在相位光栅和桥回路尚未建立或未完全稳定的情况下，即瞬态下，它们之间是有干扰的. Anderson 等描述了互泵浦相位共轭效应的稳态特性和瞬态特性并讨论了与这一效应有关的机理[23].

参见图 9.23，从 Ar^+ 激光器辐射的激光(λ=514.5 nm)，经偏振光旋转器 PR，扩束器 BX，再由分光镜 BS 分成两束，一束经反光镜 M 和透镜 L_1 射入晶体一侧(E_1)，另一束经透镜 L_2 和反光镜 M 射入晶体另一侧(E_2). 当晶体为 SBN 或 $BaTiO_3$ 时，两束光都是 e 光，输入图像 O_1，O_2 和 L_1，L_2 构成傅里叶变换系统，晶体恰恰在谱面上，所以 E_1，E_2 分别为 O_1，O_2 的傅里叶谱. 两个光路的长度不等，使 E_1，E_2 不相干，但它们的截面尺寸足够大，所以在晶体中有较大的交叠和相互作用区域. 两个输入图像前都有光束开关

S_1,S_2,相位共轭光则分别用两个分光镜 BS 引出,经中性减光板 ND 后,由两个 CCD 器件 VC_1,VC_2 探测. 分别记为 O_1^* 和 O_2^*.

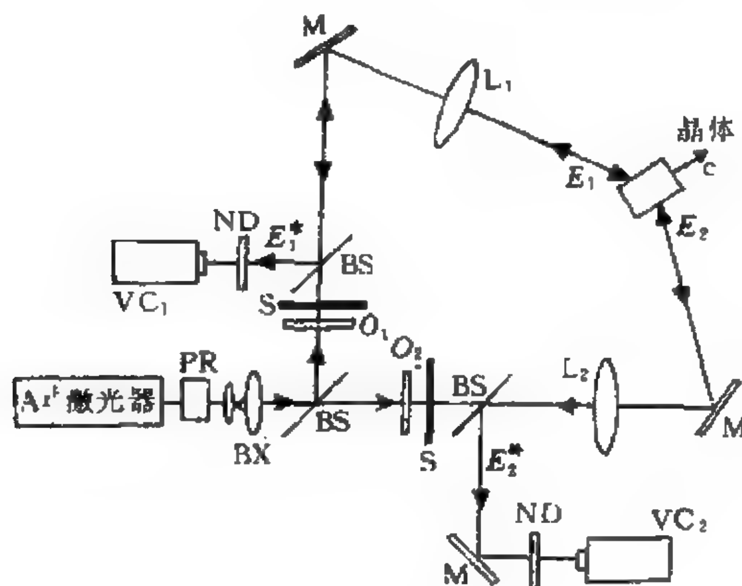


图 9.23 用 MPPC 实现图像传输

利用这一装置,可以观察到许多有趣的现象:

(1) 当一个光束,例如 E_1 突然关断后,MPPC 从稳态切变到瞬态,由它泵浦的另一光束 E_2 的相位共轭输出像 O_2^* 立即消失,但它自己的相位输出 O_1^* 却是逐渐消失的,消隐的过程对应于由 E_1 和 E_{11}建立的相位光栅 G_1 的消失过程.

(2) 当一个输入图像,例如 O_1 突然改变成 O_1', MPPC 从稳态切变到瞬态,此时 O_1' 的像立即出现在 VC_2 上,而 VC_1 上仍然是 O_1^*. 经过足够长的时间后,MPPC 回到稳态,VC_1 上逐渐出现变化后的图像,即 $O_1'^*$,而 O_2 对应的输出端 VC_2 则恢复成 O_2^*.

上面两个现象的实质都相同,即 MPPC 在瞬态下可以交换 O_1 和 O_2 的信息,而在稳态下 O_1 的输出为 O_1^*,O_2 的输出为 O_2^*.

参见图 9.24(实现装置见图 9.23). MPPC 的初始状态为稳态,两个输入图像都是高斯分布(图 9.24(a));如果突然在 O_1 处插入一个不透光的基底上的一个透明的字母 C(二元图像),MP-

PC 切变到瞬态，该图像立即出现在 VC_2 上，而 VC_1 上仍然是高斯分布(图 9.24(b))；经足够长的时间后，MPPC 回到稳态，VC_1 上出现相位共轭输出“C”，而 VC_2 仍回到高斯分布(图 9.24(c))；如果突然将 O_1 上的字母 C 移动一个小距离，MPPC 又切换到瞬态，在 VC_2 上又出现 C 的像，但 VC_1 上 C 的共轭像尚来不及变化，(图9.24(d)).

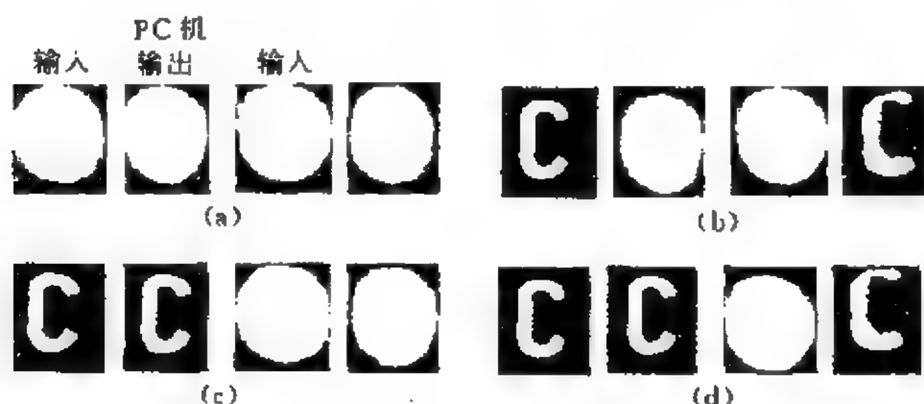


图 9.24 MPPC 的稳态和瞬态响应[28]

利用节 9.5.2 中介绍的机理，可以解释 MPPC 的瞬态响应特性. 在图 9.24(a)的情况下，E_1 和它的扇出光 E_{1f} 产生的相位光栅 G_1 为

$$\Delta n_1=\frac{n'}{2}e^{i\varphi'}\frac{A_{1f}A_1^*}{I_1}e^{-i(\boldsymbol{k}_1-\boldsymbol{k}_{1f})\cdot\boldsymbol{r}}+c.c., \tag{4}$$

E_2 和它的扇出光 E_{2f} 产生的相位光栅 G_2 为

$$\Delta n_2=\frac{n'}{2}e^{i\varphi'}\frac{A_{2f}A_2^*}{I_2}e^{-i(\boldsymbol{k}_2-\boldsymbol{k}_{2f})\cdot\boldsymbol{r}}+c.c., \tag{5}$$

运用图 9.21(a)所示的简单模型，E_{1f} 和 E_{2f} 可表为

$$E_{jf}=A_{jf}e^{i(\boldsymbol{k}_{jf}\cdot\boldsymbol{r}-\omega t)}\quad(j=1,2), \tag{6}$$

且有

$$\boldsymbol{k}_{1f}=-\boldsymbol{k}_{2f}, \tag{7}$$

由(3)式，E_{2f} 读光栅 G_1 产生 E_1 的相位共轭波

$$E_1^* \propto \Delta n_1 E_{2f} = [Cn' e^{i\varphi'} A_{1f} A_1^* e^{-i(k_1 - k_{1f})\cdot r}] \times A_{2f} e^{i(k_{2f}\cdot r - \omega t)}$$
$$= Cn' e^{i\varphi'} (A_{1f} A_{2f}) A_1^* e^{-i(k_1 \cdot r + \omega t)}, \tag{8}$$

E_{1f}读光栅 G_2 则产生 E_2 的相位共轭波

$$E_2^* \propto \Delta n_2 E_{1f} = [Cn' e^{i\varphi'} A_{2f} A_2^* e^{-i(k_2 - k_{2f})\cdot r}] \times A_{1f} e^{i(k_{1f}\cdot r - \omega t)}$$
$$= Cn' e^{i\varphi'} (A_{1f} A_{2f}) A_2^* e^{-i(k_2 \cdot r + \omega t)}, \tag{9}$$

这就是图 9.24(a)所示稳态的情形.

如果将 O_1 突然换成 O_1'(例如插入一个字母 C),E_1 立即转变成 E_1',晶体中原来的扇出光 E_{1f}立即消失,由(9)式,E_2 的相位共轭波因失去了泵浦源 A_{1f}而立即消失,VC_2 上的 O_2^* 也立即消失. 然而(8)式中 A_{1f}和 A_1^* 是由光栅 G_1(即 Δn_1)提供的,在瞬态下来不及变化,因而 E_1^* 也来不及变化. 新产生的 E_1' 可表为

$$E_1' = A_1' e^{i(k_1 \cdot r - \omega t)}, \tag{10}$$

它在 G_1 上发生衍射,衍射波可表为

$$E_1'(G_1) \propto \Delta n_1 E_1' = C'(A_{1f} A_1^*) A_1' e^{i(k_{1f} r - \omega t)}, \tag{11}$$

$E_1'(G_1)$在 G_2 再次发生衍射,最后得到

$$E_1'' = E_1'(G_1, G_2) \propto \Delta n_2 E_1'(G_1)$$
$$= C''(A_{1f} A_1^*)(A_{2f} A_2^*) A_1' e^{-i(k_2 \cdot r + \omega t)}, \tag{12}$$

由于$(A_{1f}A_1^*)$和$(A_{2f}A_2^*)$在我们的简单模型中都可以认为是高斯分布,所以 $E_1'' \propto E_1$, 再经过 L_2 的傅里叶变换,在 VC_2 出现 O_1' 的像, 见图 9.24(d). 正由于上述过程产生的是的 O_1' 像而不是 O_1' 的相位共轭像,因此其保真度不如相位共轭像. 经过足够长的弛豫时间,MPPC 回到稳态,VC_1 上又出现 $O_1'^*$,VC_2 上则又还原成 O_2^*.

9.5.4 双色泵浦光折变振荡(DCPO)

设两束输入光 E_1,E_2 的波长不相同,$\lambda_1 < \lambda_2$. 在晶体中,E_1 和它的扇出光 E_{1f}建立自己的相位光栅 G_1,E_2 和它的扇出光 E_{2f}建立自己的相位光栅 G_2. 光栅 G_1 的波矢量

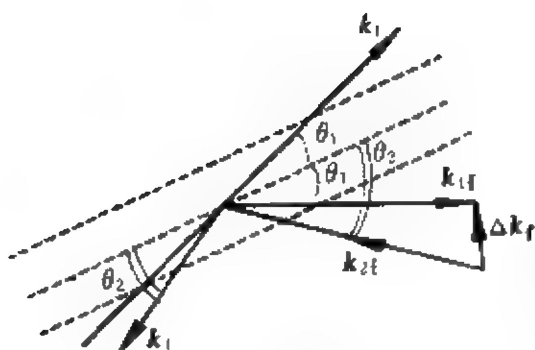


图 9.25　E_{2f}在光栅 G_1 上的衍射

$$K_1=2\pi/\Lambda_1=2k_1\sin\theta_1=4\pi\sin\theta_1/\lambda_1, \tag{13}$$

式中 θ_1 是 $\boldsymbol{k}_1$ 和 $\boldsymbol{k}_{1f}$夹角的一半，见图 9.25.

E_2 的扇出光 E_{2f}在 G_1 上发生衍射的前提，是它必须满足 Bragg 条件：

$$K_1=2\pi/\Lambda_1=2k_2\sin\theta_2=4\pi\sin\theta_2/\lambda_2, \tag{14}$$

式中 θ_2 是 $\boldsymbol{k}_{2f}$与光栅的夹角. 由式(13)、(14)得到

$$\Lambda_1=\lambda_1/2\sin\theta_1=\lambda_2/2\sin\theta_2, \tag{15}$$

由于 $\lambda_2>\lambda_1$，$\theta_2>\theta_1$，所以衍射光 E_3 的方向与 E_1 稍有不同. 当(15)式所示的 Bragg 条件关于光栅 G_1 和 G_2 同时满足时，可以同时看到输出信号 E_3 和 E_4. 事实上，在光折变晶体中，由于扇出光的角谱成分非常丰富，总有某些角谱成份的扇出光满足 Bragg 条件，形成 E_3 和 E_4. 只要 E_{1f}和 E_{2f}的方向上增益够大，振荡就能维持.

由(3)式，E_{2f}读光栅 G_1 而产生的衍射波可表为

$$\begin{aligned}E_3\propto\Delta n_1E_{2f}&=\frac{n_1^f}{2}e^{i\varphi'}\frac{A_{1f}A_1^*}{I_1}\times A_{2f}e^{-i(\boldsymbol{k}_1\cdot\boldsymbol{r}+\omega t)}e^{i\Delta\boldsymbol{k}_f\cdot\boldsymbol{r}}\\&=E_1^*e^{i\Delta\boldsymbol{k}_f\cdot\boldsymbol{r}},\end{aligned} \tag{16}$$

(16)式中

$$\Delta\boldsymbol{k}_f=\boldsymbol{k}_{2f}+\boldsymbol{k}_{1f}\neq0, \tag{17}$$

所以 E_3 不等于 E_1^*，即从晶体反射的光波不是原光波的相位共轭波. 参见图 9.25，由于

$$\Delta k_f\propto k_{1f}\Delta\theta=\pi\Delta\lambda/\lambda_1\Lambda_1\cos\theta_1, \tag{18}$$

当 $\Delta\lambda$ 很小时，(16)式中的相位畸变项 $\exp(i\Delta\boldsymbol{k}_f\cdot\boldsymbol{r})$相对于 E_1^* 而言是微扰. 由于微扰项的存在，简并消除. 然而 E_3 中依然包含 E_1 的信息. 畸变随 $\Delta\lambda$ 的增大而变大，E_3 也就越偏离 E_1^*. 参见图 9.26，$\lambda_1=488$ nm，λ_2 依次取不同的波长得到 E_3. 可以看出，当时

$\lambda_2=\lambda_1$ 时图像最清晰，此时 E_3 严格等于 E_1^*，DCPO 变成 DPC(即回到 MPPC)，$\Delta\lambda$ 越大清晰度越差. DPC 只不过是 DCPO 的特例. 这一情况和使用不同波长的激光再现全息图实像时相似.

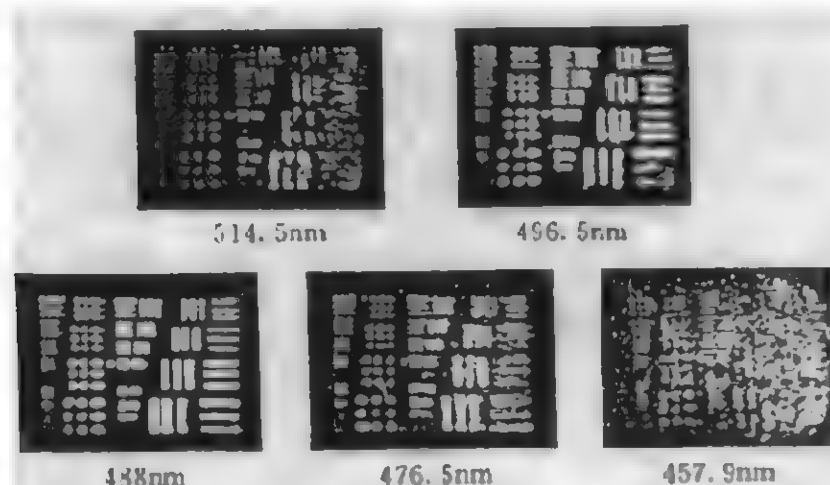


图 9.26 DCPO[16]

互泵浦相位共轭效应广泛地应用于光学信息处理中，可实现阈值控制、特殊空间光调制、边缘突出、图像相加及相减等，可参阅文献[13~24].

9.6 利用图像对相位光栅的编码实现非相干光-相干光转换

在光学信息处理的许多课题中，都需要把非相干光所荷载的图像信息转移到相干光上去，这是因为许多实用的图像信息是由非相干光产生的，而相干光的处理远比非相干光容易，所以要求实现非相干光-相干光的信息转换.

在光折变晶体中，各个光波通过相位光栅发生耦合，形成相位共轭光. 在自泵浦和互泵浦相位共轭中，相位光栅都起到了图像信息载体的作用，这一过程也可以认为是图像信息通过写入光波对相位光栅的编码. 一旦相位光栅被图像编码，其中就会包含图像

信息．当读出光读这一光栅时，输入光波中所包含的图像信息就转移到相位共轭光波中去．后一过程也可以认为是图像信息通过读出光波的解码．

Sharp 等[22]提出一个利用自泵浦相位共轭效应实现非相干光-相干光图像转换的装置，见图 9.27．一个氩离子激光器辐射的 514.5 nm 激光束经过偏振光旋转器 PR 及扩束器 BX 后，形成扩束光照射 SBN 晶体，入射光为 e 光，在 SBN 中发生自泵浦相位共轭效应．相位共轭光经分光镜 BS、透镜 L_2、反射镜 M 后成像在屏上，由相机（例如 CCD）拍摄．在没有输入图像的情况下，相位共轭光为均匀（或高斯型）光场．

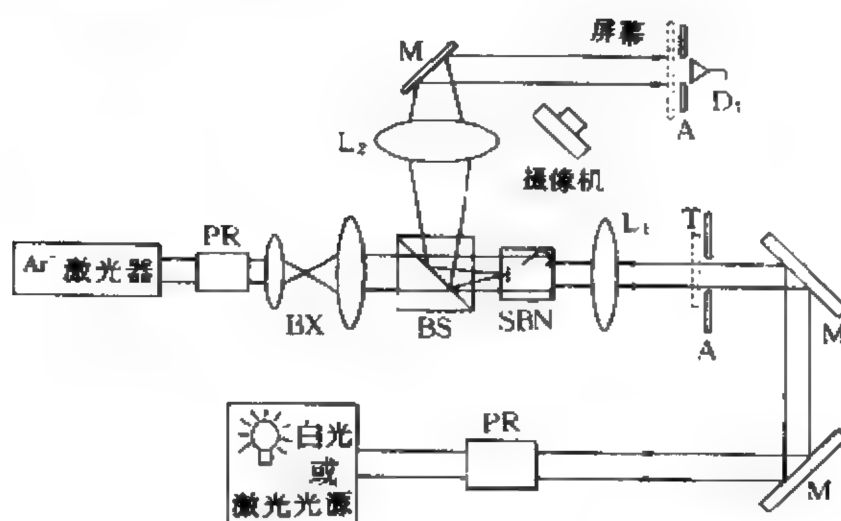


图 9.27　利用 SPPC 实现非相干光-相干光图像转换[22]

我们再用另一激光束（例如 He-Cd 激光的 441.6 nm 光束）去照射一个二元图像 T，经透镜 L_1 将 T 成像在晶体中．该写入光束为 o 光，它不参与晶体中的非线性光波耦合过程，而只是用作擦除信号，它擦除了一部分相位光栅，而且擦除的部分与 T 的像对应，这一过程实际上是对相位光栅进行负编码．当原来的入射光（e 偏振的 Ar^+ 激光）读这一经过编码手续的光栅时，反射的激光束中就出现了 T 的负像．显然写入光波不一定是激光，用非相干

光(氙灯)经过光阑 A 变成细光束代替 He-Cd 激光作为写入光，同样在反射的激光束中出现 T 的负像. 这样就实现了图像信息从非相干光向相干光的转换. 在这一实验中，写入光的强度、位置都要经过仔细的调节，写入光的强度不宜过强，否则会过度破坏相位光栅. 图 9.28(a)为输入的非相干图像(分辨率板)，(b)为相干输出图像，分辨率约 28 lp/mm.

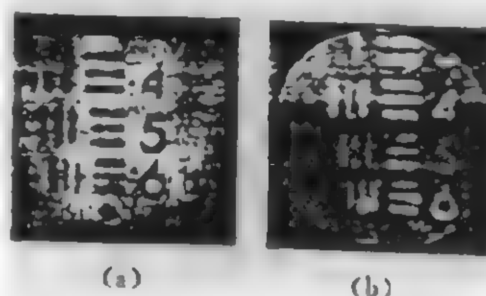


图 9.28 非相干光-相干光图像转换结果[22]
(a)非相干输入图像 (b)相干输出图像

利用光折变晶体中的四波混频效应同样可以实现上述转换. 图 9.29 表示这样一个系统，它是由 Shi 等提出的[19]. 所用的晶体为 BSO，它在写入光 I_1，I_2 和读出光 I_3 的作用下产生四波混频效应，生成相位共轭光 I_4. 用一个非相干光照射图像 T，并投射到晶体中，通过擦除效应对相位光栅进行调制或编码，图像信息就经过分光镜 BS 转移到输出的 I_4 中. 该实验的输出相干光图像分辨率达到 14.3 lp/mm[19]. 图中用氙灯作为光源 S，F 为 545 nm 滤光镜，透镜 L_1 和 L_2 构成成像系统.

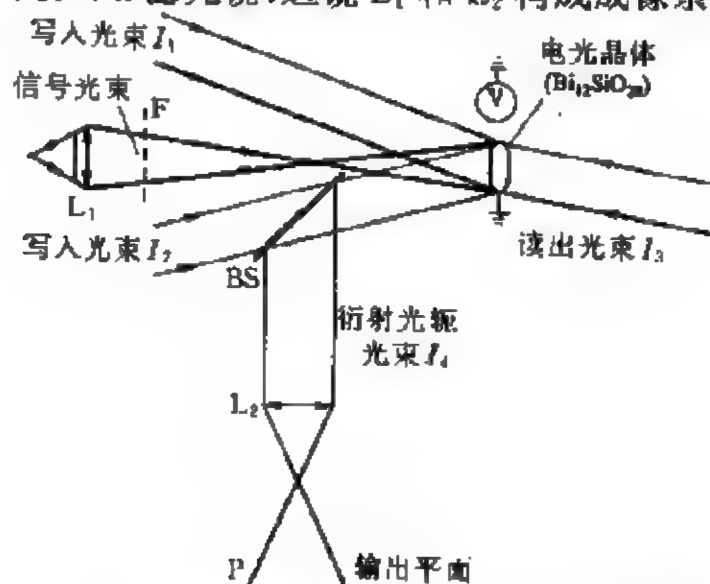


图 9.29 利用四波混频实现非相干光-相干光图像转换[19]

9.7 相位共轭在联合傅里叶变换相关识别中的应用

光折变效应能产生高保真的相位共轭像，并具有很高的效率，几十 mW 或几 mW 的输入就能获得足够强的相位共轭光输出，因此大大拓宽了它的应用. 本节讨论光折变效应在联合傅里叶变换相关(JTC)中的应用.

JTC 是图形识别的重要方法. 但是在许多重要的图像处理课题中，要求输入物体本身不能有相位畸变. 然而，作为光电混合相关识别系统的重要输入器件——空间光调制器(SLM)通常都会有附加的相位畸变，它严重影响了 JTC 的识别效率. 尽管可以用液门来部分补偿相位畸变，但它原则上只能补偿外部的相位畸变，而无法补偿 SLM 的内部相位畸变.

JTC 的另一个问题是相关光斑的尺寸过大，且互相关斑过亮，影响了识别效果. 为了提高识别的准确度，一个行之有效的办法是对物体进行预先编码(pre-encoding). 而相位共轭的非线性很适合于这一手续. 这一方案是 Yu 和 Jutamulia[7]提出的. 参见图 9.30. 一对待识别的图形 O_1 和 O_2 由 CCD_1 探测后，经计算机及空间光调制器 SLM 输入系统. 由 Ar^+ 激光器辐射的激光(514.5 nm)分成三路，一路通过分光镜 BS_1，反光镜 M_2 照射 BSO 晶体，成为参考光；第二路经 BS_2，M_3 及 BS_3 照射输入物体 O_1 和 O_2，再通过透镜 L_1 成像到 BSO 晶体上，成为物光. 物光和参考光对晶体而言是写入光，它们相干叠加形成相位光栅. 第三束光经 M_1，M_5 和 M_4 后，沿着与参考光相反的方向照射晶体，成为读出光. 光路的安排使读出光的长度超出写入光的相干长度范围，这样一来，读出光与写入光不能形成光栅，系统构成二波混频机制. 读出光在相位光栅上的衍射光束构成物光的相位共轭光束，它准确地沿着物光的相反方向，经透镜 L_1 成像到 SLM 上，再经 BS_3 分

光，通过透镜 L_2 的傅里叶变换，在输出平面 P_0 上形成物体的相关输出．

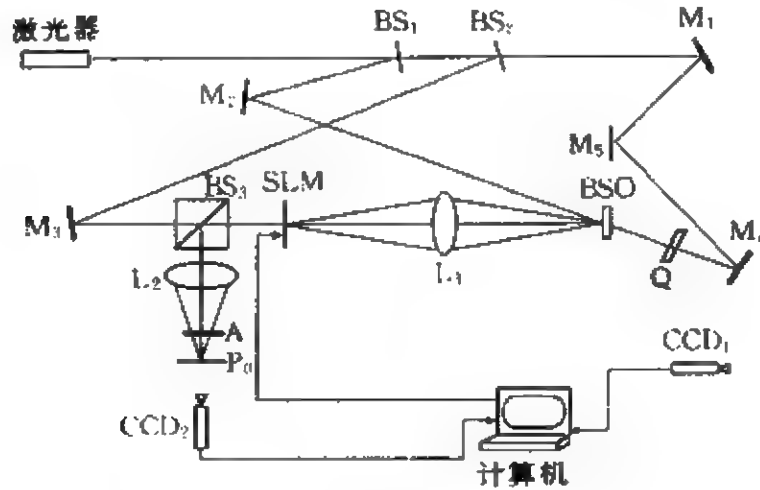


图 9.30　光折变相位共轭 JTC[17]

· ·对波片 Q，A 用来消除杂散光，提高信噪比．设物光为

$$O(x,y)\exp[i\varphi(x,y)]=O_1(x-b,y)\exp[i\varphi_1(x-b,y)] \\ +O_2(x-b,y)\exp[i\varphi_2(x+b,y)], \quad (1)$$

式中 O_1，O_2 是两个物体的振幅，φ_1，φ_2 分别是附加相位，$2b$ 为两个物体的平均间距，设 O_1，O_2 为正实数．在像平面(BSO 晶体)上物光可表为

$$O\left(\frac{x}{M},\frac{y}{M}\right)\exp\left[i\varphi\left(\frac{x}{M},\frac{y}{M}\right)\right]\exp\left[ik\left(\frac{x^2+y^2}{2Mf}\right)\right], \quad (2)$$

式中 M 为放大率，f 是 L_1 的焦距．从晶体上反射的光波可表为

$$O'\left(\frac{x}{M},\frac{y}{M}\right)\exp\left[-i\varphi\left(\frac{x}{M},\frac{y}{M}\right)\right]\exp\left[-ik\left(\frac{x^2+y^2}{2Mf}\right)\right] \\ \times\exp\left[i\theta\left(\frac{x}{M},\frac{y}{M}\right)\right], \quad (3)$$

严格地说，它并不是物光准确的相位共轭光波，O' 和 O 的差别是由晶体的非线性造成的．此外还有一个相位项 $\exp(i\theta)$，如在晶体上没有外加电压，则 $\theta=90°$．

经 L_1 成像到空间光调制器 SLM 上的光波为

$$O'(x,y)\exp[-i\varphi(x,y)]\exp[i\theta(x,y)], \quad (4)$$

再次通过 SLM 后，恰恰抵消了 SLM 引起的相位畸变 $\varphi(x,y)$：

$$O(x,y)O'(x,y)\exp[\mathrm{i}\theta(x,y)]. \tag{5}$$

这一手续又称相位补偿. 由于相位畸变项已完全消除，大大提高了识别效果.

图 9.31(a)是一对待识别的图形，背景中有随机噪声. 图(b)是它们的 JTC 谱. 图(c)是经过相位补偿后的 JTC 谱. 可以明显地看出，相位补偿已大大抑制了 Joint 变换的噪声. 图 9.32 是它们的相关输出，图(a)是不带相位补偿的情况，图(b)是带有相位补偿的情况. 在图(a)中两个相关峰淹没在 0 级谱项和本底噪声中，几乎分辨不出来；在图(b)中噪声本底大大减小，0 级谱项强度减弱，范围缩小，从而两个相关峰很清晰. 我们看到，通过相位补偿手续，JTC 相关输出的信噪比大大提高.

至于(5)式中出现的相位因子 $\exp[\mathrm{i}\theta(x,y)]$，与物光光强、物光参考光强度比有关，可以用来对物体进行"预先编码"(pre-encoding)，处置恰当，有助于提高识别率.



图 9.31　JTC 谱[7]

(a) 带有随机噪声的输入图形

(b) 未经相位补偿的 JTC 谱

(c) 经过相位补偿的 JTC 谱

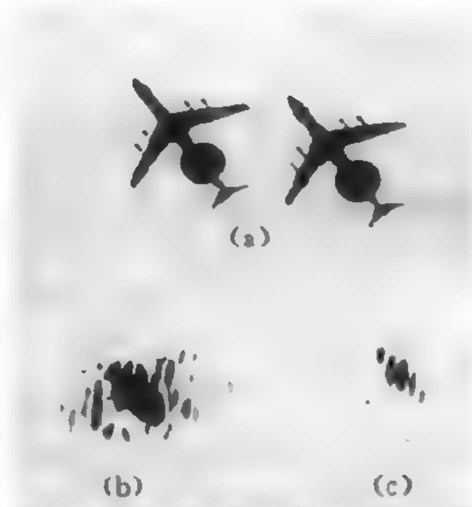


图 9.32　重建的相关输出[7]

(a)未经相位补偿的相关光斑

(b)经过相位补偿的相关光斑

9.8 光折变非线性联合变换相关器

过去设计的各种形式的光折变联合傅里叶变换相关识别系统(JTC),如果用峰值-噪声比(PNR)或用鉴别能力(即鉴别不同于参考信号的任何其他信号的能力)来衡量,都比不上纯相位滤波相关识别器(POF)[27,36,37]. 究其原因,这些光折变JTC的响应特性仍然与匹配滤波很相似,因此与匹配滤波有同样的缺点:当噪声具有高斯型统计规律时,判别效率是令人满意的,但当本底噪声比较杂乱,不符合高斯统计时,匹配识别的效果就比较差. 实际工作中要求识别的信号一般都混在本底噪声之中,而噪声的情况比较复杂,许多噪声都不符合高斯统计,这样一来识别的效果就不能令人满意.

一个有效的处理方案,是引入非线性探测来改进谱面的响应特性,例如取阈值,或使响应曲线具有趋向于饱和的特征等,这些非线性的处理手续可以增大信号中高频成分的权重.

我们知道,JTC是一个三步的过程:第一步是对一对(或多个)待识别的图形进行第一次傅里叶变换,得到它们的谱;第二步是谱面的平方律转换,把复振幅分布转换成功率谱;最后一步是第二次傅里叶变换,把功率谱转换成相关输出. 因此所有的JTC系统都有一个谱面的平方律探测器,它可以是空间光调制器,也可以是CCD再加上计算机数字图像处理器. 这些中介的电子器件使处理速度减缓,使系统失去了光学处理的高度并行的特性. 所以JTC的设计者都希望找到一个纯光学的平方律转换器件.

最近,Khoury等[25,27]提出一种光折变二波混频非线性JTC. 除最后的相关输出平面的监视采用CCD以外,其他器件都是光学器件,发挥了处理系统自身的并行特性;其次,他们利用了光折变效应的非线性,使系统对大信号的响应特性接近饱和(平方根律),从而得到了可与纯相位匹配滤波识别(POF)媲美的识别效果.

图 9.33 即全光学光折变 JTC 实验系统示意图. S 为待识别的一对图形，该图形记为 g，可以是透明基底上的真实图像，也可以是用空间光调制器 SLM 输入的图像，用光束 E_1 照明，再通过 L_1 进行傅里叶变换，谱面上放置光折变晶体 PRC($BaTiO_3$)，作为物光；参考光 E_4 经扩束镜 BE 后照射 PRC，与物光进行二波混频，产生相位光栅，输入图形的傅里叶谱就记录在相位光栅中. 物光穿过 PRC，经过 L_2、自泵浦相位共轭反射镜 SPPC，再反向通过 L_2，射入 PRC，构成 1∶1 成像系统. SPPC 的自对准特性，保证射入 PRC 的 E_2 与 E_1 方向相反，严格符合 Bragg 条件，所以在相位光栅上衍射产生 E_3，这是一个四波混频过程，E_2 读相位光栅产生 E_3，E_3 就是 E_1 的相位共轭光 E_1^*，它与探测光（参考光）的方向相反，在分光镜 BS 上反射，并由 L_3 作最后一次傅里叶变换，输出的相关信号由 CCD 监视.

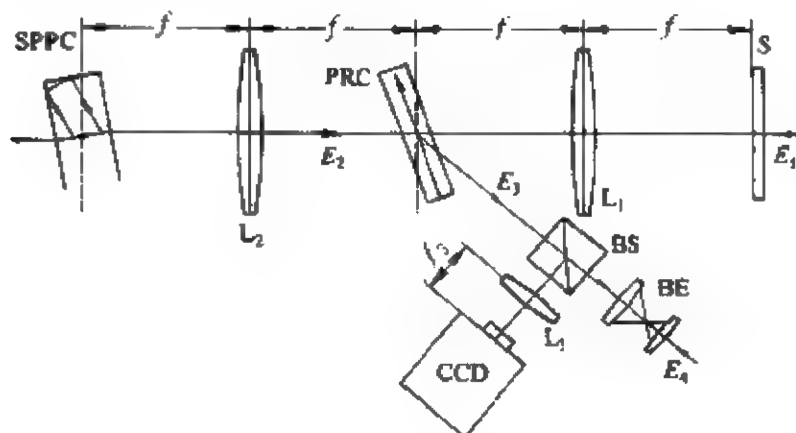


图 9.33 全光学光折变 JTC 实验系统[25]

由节 9.3 四波混频的耦合波方程(28)第三式，如忽略晶体的吸收，并设 $|A_3| \ll |A_1|$，$|A_2|$，则该式可近似表为

$$A_3 = \frac{\gamma L}{2I_0} A_1' A_4^* A_2', \tag{1}$$

式中 γ 是复耦合系数，$\gamma = -\Gamma/2$，采用的是文献[1]中的符号，L 是晶体中的作用长度，I_0 是总光强，

$$I_0=I_1+I_2+I_4, \tag{2}$$

A_1' 是物光的复振幅，可表为 g 的傅里叶谱和照射光 E_1 的振幅 $|A_1|$ 的积：

$$A_1'=\frac{|A_1|}{\lambda f_1}G\exp(i\varphi_1), \tag{3}$$

其中 λ 是波长，f_1 是 L_1 的焦距，G 和 φ_1 分别是 g 的傅里叶谱的振幅和相位函数.

E_1 透过 PRC，并经 L_1，SPPC 再经 L_1 折回的光波是 PRC 相位光栅的读出光. 这一循环从物理上看不过是 1：1 的成像过程，所以该读出光仍然是 g 的傅里叶谱，它的复振幅

$$A_2'=\sqrt{\beta}\frac{|A_1|}{\lambda f_1}G^*\exp(i\varphi_2), \tag{4}$$

式中 β 为上述循环的综合反射系数，φ_2 则是 E_2 的相位. 由于 SPPC 的自泵浦相位共轭效应，在反射光中出现 G 的复共轭 G^*. 以(2)，(3)及(4)式代入(1)，得到

$$A_3=f(G)=\frac{\gamma L\sqrt{\beta}(\lambda L)^{-2}I_1A_4^*|G|^2\exp[i(\varphi_1+\varphi_2)]}{(\beta+1)(\lambda f)^{-2}I_1|G|^2+I_4} \tag{5}$$

上式还可以表为

$$f(G)=Z_c\frac{m_e|G|^2}{1+m_e|G|^2}, \tag{6}$$

式中

$$Z_c=\frac{\gamma L\sqrt{\beta}\{A_4^*\exp[i(\varphi_1+\varphi_2)]\}}{\beta+1} \tag{7}$$

(6)式中等效光强比

$$m_e=\frac{I_1(\beta+1)}{(\lambda f)^2I_4}, \tag{8}$$

(5)式或(6)式表示系统输出的相位共轭信号 A_3 作为输入信号 g 的傅里叶频谱 G 的函数，可以看作是系统对输入信号的响应函数. 我们就从 $f(G)$ 入手来研究系统的响应特性.

当信号较小时，$\sqrt{m_e}G \ll 1$，则由(5)式，有

$$f(G) \propto Z_c m_e |G|^2, \tag{9}$$

响应特性符合平方律，这正是一般 JTC 的要求.

当信号较大时，$\sqrt{m_e}G \gg 1$，则有

$$f(G) \propto Z_c \left(1 - \frac{1}{m_e |G|^2} \right) \approx Z_c, \tag{10}$$

表示系统对大信号的响应特性趋向于饱和，而在平方律和饱和区段之间，则趋近于平方根律，这正是纯相位匹配滤波(POF)的响应特性，因此系统对大信号的响应特性和 POF 很相似. 图 9.34 是本系统的响应曲线. 图 9.35 给出计算机模拟的相关输出结果，其中图(a)(c)(e)为相似图形的相关输出，图(b)(d)(f)为不相似图形的相关输出[25]. 图(a)(b)是平方律区间的相关输出，图(c)(d)是过渡区间，图(e)(f)是饱和区间的相关输出，可以看出相关峰非常尖锐，具有纯相位型匹配滤器的效果.

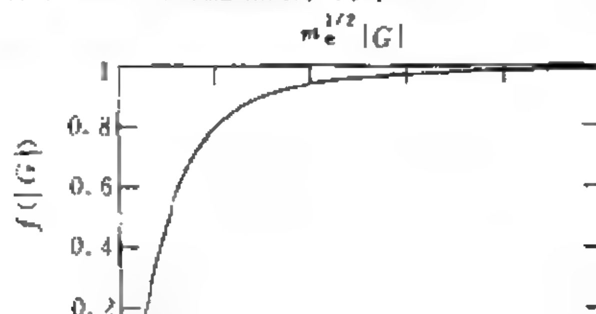


图 9.34　光折变效应非线性联合变换相关器(JTC)的响应曲线

综上所述，本系统的响应特性是“软截止”(soft clipping)而不是“硬截止”(hard clipping). 硬截止相当于取阈值处理. 阈值的大小是非常敏感的，阈值过低，许多噪声谱可能超过阈值形成误判；阈值过高，则一些信噪比不够高的信号可能淹没在噪声本底中测不出来形成漏判.

具有软截止响应特性的 JTC 系统，其性能比较优越，而光折变效应的高度非线性正好可以利用，设计成功在不同信号强度区间具有不同响应特性的相关器.

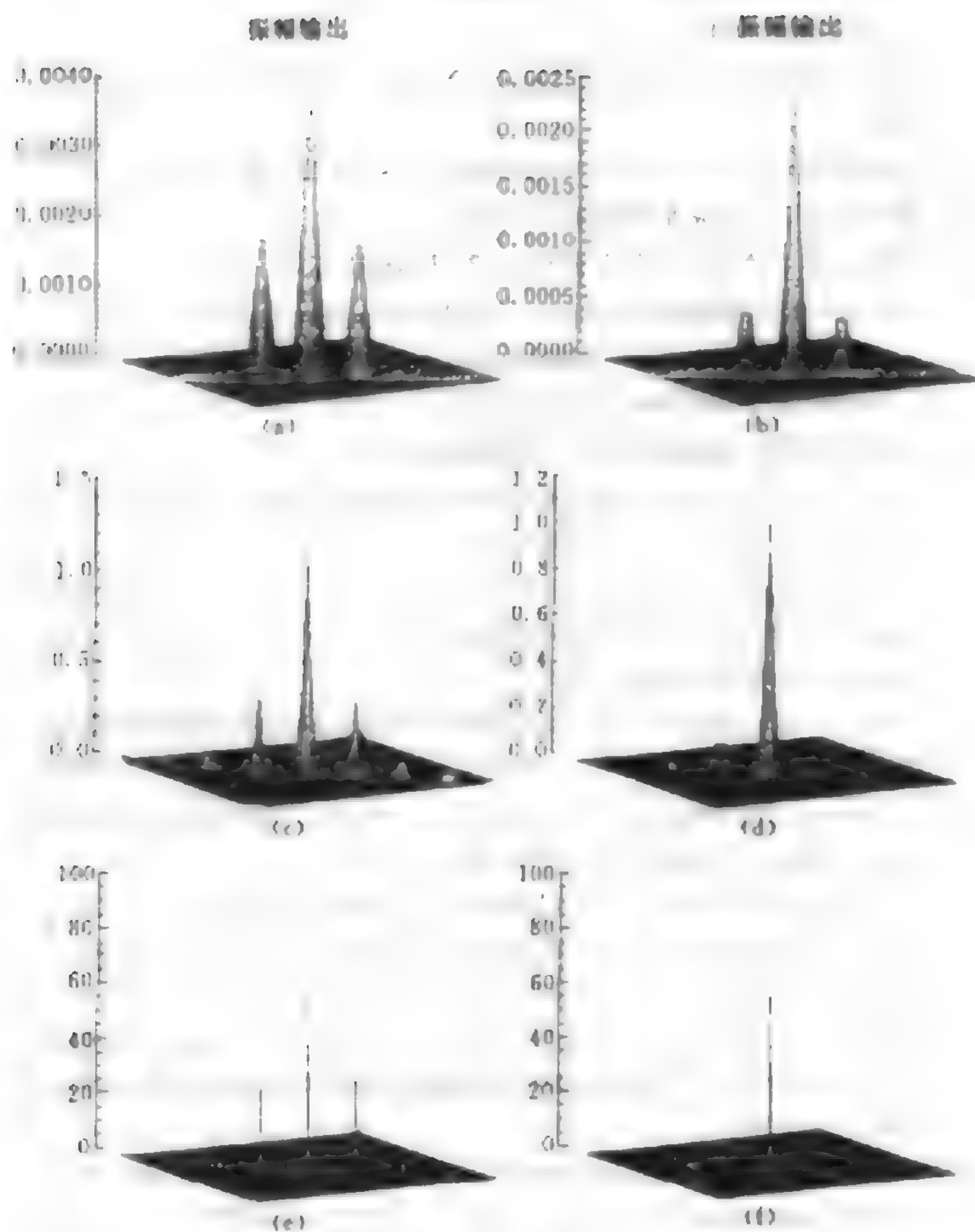


图 9.35

参考文献

[1] M. Cronin-Golomb, B. Fischer, J. O. White, and A. Yariv, Theory and applic a

tions of four-wave mixing in photorefractive media, *IEEE. J. Quantum Electron.*, **QE-20**(1984), 12-29.

[2] P. Yeh, Introduction to Photorefractive Nonlinear Optics, Wiley, New York, 1993.

[3] P. Yeh, Photorefractive phase conjugators, *Proc. IEEE*, **80**(1992), 435-450.

[4] 曹昌祺,电动力学,高等教育出版社,1962.

[5] 刘思敏,郭儒,凌振芳,光折变非线性光学,中国标准出版社,1992.

[6] P. Yeh, Optical Waves in Crystals, John Wiley & Sons, New York, 1984.

[7] F. T. S. Yu and S. Jutamulia, Optical Signal Processing, Computing, and Neural Networks, John Wiley & Sons, New York, 1992.

[8] 羊国光,宋菲君,高等物理光学,中国科技大学出版社, 1991.

[9] J. Feinberg, Self-pumped, continuous-wave phase conjugator using internal reflection, *Opt. Lett.*, **7** (1982), 486-488

[10] Y. Zhu, C. Yang, M. Hui, X. Niu, J. Zhang, T. Zhou, and X. Wu, Phase conjugation of $BaTiO_3$:Ce by backward stimulated photorefractive scattering, *Appl. Phys. Lett.*, **64**(1994), 2341-2343.

[11] Y. Zhu, X. Wu, and G. Yang, Spectroscopic and self-pumped phase conjugation of visible-sensitive Cerium-doped Barium Titenate, SPIE, **2896**(1996), 69-77.

[12] J. Feinberg, Asymmetric self-defocusing of an optical beam from the photorefractive effect, *J. Opt. Soc. Am.*, **72**(1982), 46-51.

[13] E. J. Sharp, W. W. Clar III, M. J. Miller, G. L. Wood, B. Monson, G. J. Salamo, and R. R. Neurgaonkar, Double phase conjugation in a tungsten bronze crystals, *Appl. Opt.*, **29**(1990), 743-749.

[14] S. W. James, and R. W. Eason, Intensity-dependent thresholding and switching in the photorefractive bridge mutually pumped phase conjugator, *Opt. Lett.*, **16** (1991), 551-553.

[15] S. Sternklar, S. Weiss, and B. Fischer, Optical information processing with the double phase conjugate mirror, *Opt. Eng.*, **26**(1987), 423-427.

[16] S. Sternklar, and B. Fischer, Double-color-pumped photorehractivre oscillator and image color conversion, *Opt. Lett.*, **12**(1987), 711-713.

[17] E. Ochoa, L. Hesselink, and J. W. Goodman, Real-time intensity inversiopn using two-wave and four-wave mixing in photorefrective $Bi_{12}GeO_{20}$, *Appl. Opt.*, **24**(1985), 1826-1832.

[18] S. G. Rabbani, J. L. Shultz, G. J. Salamo, E. J. Sharp, W. W. Clark III, M. J. Miller, G. L. Wood, and R. R. Neurgaonkar, Color Imaging in Photorefractive Crystals, *Appl. Phys.*, **B53**(1991), 323-329.

[19] Y. Shi, D. Psaltis, A. Marrakchi, and A. R. Tanguay, Photorefractive incoherent-to-coherent optical converter, *Appl. Opt.*, **22**(1983), 3665-3667.

[20] M. B. Klein, G. J. Dunning, G. C. Valley, R. C. Lind, and T. R. O'Meara, Imageing threshold detector using a phase-conjugate resonator in BaTiO3, *Opt. Lett.*, **11**(1986), 575-577.

[21] J. Ma, L. Liu, S. Wu, and Z. Wang, Grating-encoded multichannel photorefractive incoherent-to-coherent optical conversion, *Opt. Lett.*, **14** (1989), 572-574.

[22] E. J. Sharp, J. L. Wood, W. W. Clark III, G. J. Salamo, and R. R. Neurgaonkar, Incoherent-to-coherent conversion using a photorefractive self-pumped phase conjugator, *Opt. Lett.*, **17**(1992), 207-209.

[23] R. J. Anderson, E. J. Sharp, G. L. Wood, and G. J. Salamo, Image transfer by mutually pumped phase conjugators, *Appl. Opt.*, **35**(1996), 854-859.

[24] P. Yeh, T. Y. Chang, and M. D. Ewbank, Model for mutually pumped phase conjugation, *J. Opt. Soc. Am.*, **B5**(1988), 1743-1749.

[25] J. Khoury, J. S. Kane, G. Asimellis, M. Cronin-Golomb, and C. Woods, All-optical nonlinear joint Fourier transform correlator, *Appl. Opt.*, **10**(1994), 8216-8225.

[26] S. Granieri, M. d. C. Lasprilla, N. Bolognini, and E. E. Sicre, Space variant optical correlator based on the fractional Fourier transform: implementation by the use of a photorefractive $Bi_{12}GeO_{20}$ (BGO) holographic filter, *Appl. Opt.*, **35** (1996), 6951-6954.

[27] J. Khoury, M. Cronin-Golomb, P. Gianino, and C. Woods, Photorefractive two-beam-coupling nonlinear joint-transform correlator, *J. Opt. Soc. Am.*, **B11** (1994), 2167-2174.

[28] K. R. McDonald and J. Feinberg, Theary of a self-pumped phase conjugator with two coupled interaction ragions, *J. Opt. Soc. Am.*, **B73**(1983), 548-553.

[29] Y. Lian, H. Gao, P. X. Ye, Q. Guan, and J. Wang, Self-pumped phase conjugation with a new mechanism in $KTa_{1-x}Nb_xO_3$ crystals, *Appl. Phys. Lett.*, **63** (1993), 1745-1747.

[30] Y. Lian, S. X. Dou, H. Gao, Y. Zhu, X. Wu, C. Yang, and P. X. Ye, Mechanism transformation with wavelength of self- pumped phase conjugation in BaTiO3:Ce, *Opt. Lett.*, **19**(1994), 610-612.

[31] Y. Lian, H. Gao, S. X. Dou, H. Wang, P. X. Ye, Q. Guan, and J. Wang, Mechanism transition of self-pumped phase conjugation in $KTa_{1-x}Nb_xO_3$:Fe crystals, *Appl. Phys.*, **B59**(1994), 655-658.

[32] Y. Lian, S. X. Dou, J. Zhang, H. Gao, Y. Zhu, X. Wu, C. Yang, and P. X. Ye, Variation of mechanism transition wavelength of self-pumped phase conjugation with Ce-content in $BaTiO_3$:Ce crystals, *Opt. Commun.*, **110**(1994), 192-196.

[33] S. X. Dou, H. Gao, J. Zhang, Y. Lian, H. Wang, Y. Zhu, X. Wu, C. Yang, and P. X. Ye, Sdudies on formation mechanisms of self-pumped phase conjugation in $BaTiO_3$:Ce crystals at wavelength from 570 to 680 nm, *J. Opt. Soc. Am.*, **B12** (1995), 1048-1055.

[34] S. X. Dou, J. Zhang, M. G. Wang, H. Gao, and P. X. Ye, Theorectal study on effects of stimulated phtorefractive backscattering in self-pumped phase conjugators, *J. Opt. Soc. Am.*, **B12**(1995), 1056-1064.

[35] P. X. Ye, What factors determine the generation mechanism of photorefractive self-pumped phase conjugation, Proc. SPIE, **2896**(1996), 62-67.

[36] J. L. Horner and P. D. Gianino, Phase-only matched filtering, *Appl. Opt.*, **23** (1984), 812-816.

[37] J. L. Horner and J. R. Leger, Pattern recognition with binary phase-only filters, *Appl. Opt.*, **24**(1985), 609-611.

[38] T. Y. Chang and R. W. Hellwarth, Optical phase conjugation by backscattring in barium titanate, *Opt. Lett.*, **10**(1985), 408-410.

[39] J. E. Ford; Y. Fainman; and S. H. Lee, Enhancedphotorefractive performance from 45(-cut $BaTiO_3$, *Appl. Opt.*, **28**(1989), 4808-4815.

[40] R. A. Vazquez; F R. Vachss, R R. Neurgaonkar, and M. D. Ewbank, Large photorefractive coupling coefficient in a thin cerium-doped strontium barium niobate crystal, *J. Opt. Soc. Am.*, **B8** (1991), 1932-1941.

[41] W. S. Rabinovich; B. J. Feldman, and G. C. Gilbreath, Suppression of photorefractive beam fanning using achromatic gratings, *Opt. Lett.*, 16(1991), 1147-1149.

[42] H. Rajbenbach, A. Delboulbe, and J. P. Huignard, Noise suppression in photorefractive image amplifiers, *Opt. Lett.*, **14**(1989), 1275-1277.

[43] G. C. Valley, Competition between forward- and backward-stimulated-photorefractive scattering in $BaTiO_3$, *J. Opt. Soc. Am.*, **B4**(1987), 14-19.

[44] C. Gu, and P. Yeh, Scattering due to randomly distributed charge particlesin photorefractive crystals, *Opt. Lett.*, **16**(1991), 1572-1574.

[45] Y. H. Hong, P. Xie, J. H. Dai, Y. Zhu, H. G. Yang, and H. Zhang, Fanningeffects in photorefractive crystals, *Opt. Lett.*, **18**(1993), 772-774.

[46] P. Xie, Y. Hong, J. Dai, Y. Zhu, and H. Zhang, Theoretical and experimentalstudies of fanning effects in photorefractive crystals, *J. Appl. Phys.*, **75** (1993), 813-818.

[47] P. Xie, J. Dai, P. Wang, and H. Zhang, A two-dimensional theory of beamfanning in photorefractive crystals, *J. Appl. Phys.*, **74**(1994), 1891-1895.

[48] P. Xie, J. Dai, P. Wang, and H. Zhang, Spetial fidelity of imageamplification in photorefractive crystals, *Appl. Opt.*, **35**(1996), 7102-7107.

[49] P. Xie, J. Dai, P. Wang, and H. Zhang, Self-pumped conjugation inphotorefractive crystals: reflectivity and special fidelity, *Phys. Rev.*, **A55**(1997), 3092-3100.

附录 A9　耦合波方程的解

A9.1　二波混频方程的解

节 9.3 二波混频的耦合波方程为

$$\frac{\mathrm{d}A_1}{\mathrm{d}z}=-\frac{\Gamma}{2I_0}A_2A_2^*A_1-\frac{\alpha}{2}A_1,\tag{1}$$

$$\frac{\mathrm{d}A_2}{\mathrm{d}z}=\frac{\Gamma}{2I_0}A_1A_1^*A_2-\frac{\alpha}{2}A_2,\tag{2}$$

其中 Γ 为复耦合系数,

$$\Gamma=\gamma+\mathrm{i}2\beta,\tag{3}$$

γ 和 β 都是实数. 定义

$$\Gamma/2I_0=\mu+\mathrm{i}\nu,\tag{4}$$

这样定义的 μ 和 γ 有如下关系:

$$\mu=\gamma/2I_0.\tag{5}$$

我们再设

$$A_1=\rho_1\mathrm{e}^{\mathrm{i}\theta_1},\tag{6}$$

$$A_2=\rho_2\mathrm{e}^{\mathrm{i}\theta_2},\tag{7}$$

ρ_1,ρ_2,θ_1 和 θ_2 均为实函数. 原方程(1),(2)分别用 A_1,A_2 除,再将(4)、(6)及(7)代入,得到

$$\frac{\mathrm{d}}{\mathrm{d}z}(\ln\rho_1+\mathrm{i}\theta_1)=-(\mu+\mathrm{i}\nu)\rho_2^2-\frac{\alpha}{2},\tag{8}$$

$$\frac{\mathrm{d}}{\mathrm{d}z}(\ln\rho_2+\mathrm{i}\theta_2)=(\mu+\mathrm{i}\nu)\rho_1^2-\frac{\alpha}{2},\tag{9}$$

将方程(8)、(9)的实部和虚部分开,其中关于 ρ 的方程为

$$\frac{\mathrm{d}}{\mathrm{d}z}\ln\rho_1=-\mu\rho_2^2-\frac{\alpha}{2},\tag{10}$$

$$\frac{\mathrm{d}}{\mathrm{d}z}\ln\rho_2=\mu\rho_1^2-\frac{\alpha}{2},\tag{11}$$

(10)、(11)可以改写为

$$\rho_1 \frac{\mathrm{d}}{\mathrm{d}z}\rho_1 = -\mu\rho_1^2\rho_2^2 - \frac{\alpha}{2}\rho_1^2, \tag{12}$$

$$\rho_2 \frac{\mathrm{d}}{\mathrm{d}z}\rho_2 = \mu\rho_1^2\rho_2^2 - \frac{\alpha}{2}\rho_2^2, \tag{13}$$

两式相加并积分，得

$$\rho_1^2 + \rho_2^2 = C_1 \mathrm{e}^{-\alpha z}, \tag{14}$$

C_1 为常数. (10)式和(11)相减，代入(14)的结果并求积分，得到

$$\ln\rho_1 - \ln\rho_2 = -\mu C_1 \mathrm{e}^{-\alpha z},$$

即

$$\frac{\rho_1}{\rho_2} = C_2 \exp\left(\frac{\mu C_1}{\alpha}\mathrm{e}^{-\alpha z}\right). \tag{15}$$

将(15)代入(14)，得到

$$I_2 = \rho_2^2 = \frac{C_1 \mathrm{e}^{-\alpha z}}{1 + C_2^2 \exp\left(\frac{\mu C_1 \mathrm{e}^{-\alpha z}}{\alpha}\right)}. \tag{16}$$

以(16)代入(15)，得到

$$I_1 = \rho_1^2 = \frac{C_1 \mathrm{e}^{-\alpha z}}{1 + C_2^{-2} \exp\left(-\frac{\mu C_1}{\alpha}\mathrm{e}^{-\alpha z}\right)}. \tag{17}$$

设 $I_1(0)$和 $I_2(0)$为入射面上的光强，则有

$$I_2(0) = \frac{C_1}{1 + C_2^2 \exp\left(\frac{2\mu C_1}{\alpha}\right)}, \tag{18}$$

$$I_1(0) = I_2(0) C_2^2 \exp\left(\frac{2\mu C_1}{\alpha}\right). \tag{19}$$

引入光强比

$$m = \frac{I_1(0)}{I_2(0)}, \tag{20}$$

从(18)、(19)得到

$$C_2^2 \exp\left(\frac{2\mu C_1}{\alpha}\right) = m, \tag{21}$$

$$C_1=(1+m)I_2(0). \tag{22}$$

在弱吸收近似下：

$$C_2^2\exp\left(\frac{2C_1\mu}{\alpha}e^{-\alpha z}\right)\approx C_2^2\exp\left(\frac{2C_1\mu}{\alpha}\right)\exp(-2C_1\mu z)$$

$$=m\exp(-\gamma z), \tag{23}$$

其中用到(5)及(22)式，最后

$$I_1(z)=\frac{1+m^{-1}}{1+m^{-1}e^{\gamma z}}e^{-\alpha z}, \tag{24}$$

$$I_2(z)=\frac{1+m}{1+me^{-\gamma z}}e^{-\alpha z}, \tag{25}$$

即节 9.3(18)式.

A9.2 四波混频方程的解

四波混频的耦合波方程(节 9.3(28)式)在未耗尽近似下可表为

$$\frac{dA_1}{dz}=-\frac{\alpha}{2}A_1, \tag{26}$$

$$\frac{dA_2^*}{dz}=\frac{\alpha}{2}A_2^*, \tag{27}$$

$$\frac{dA_3}{dz}=-\frac{\Gamma}{2I_0}(A_1A_4^*+A_2^*A_3)A_2+\frac{\alpha}{2}A_3, \tag{28}$$

$$\frac{dA_4^*}{dz}=-\frac{\Gamma}{2I_0}(A_1A_4^*+A_2^*A_3)A_1^*-\frac{\alpha}{2}A_4^*. \tag{29}$$

(26)、(27)式可直接解出：

$$A_1(z)=A_1(0)e^{-\alpha z/2}, \tag{30}$$

$$A_2(z)=A_2(L)e^{\alpha(z-L)/2}, \tag{31}$$

容易证明

$$\frac{d}{dz}(A_1^*A_3-A_2A_4^*)=0, \tag{32}$$

所以有守恒量

$$A_1^* A_3 - A_2 A_4^* = C_0. \tag{33}$$

其中 C_0 为常数. (28)式可利用(33)式改写为

$$\frac{\mathrm{d}A_3}{\mathrm{d}z} - \frac{1}{2}(\alpha - \Gamma)A_3 = \frac{C_0 \Gamma}{2I_0} A_1, \tag{34}$$

式中 A_1 由(30)式表示，I_0 则为

$$I_0 \approx I_1 + I_2 = A_1(0)A_1^*(0)\mathrm{e}^{-\alpha z}(1 + r\mathrm{e}^{2\alpha z - \alpha L}), \tag{35}$$

其中 r 为泵浦光强度比：

$$r = \left|\frac{A_2(L)}{A_1(0)}\right|^2 = \frac{I_2(L)}{I_1(0)}, \tag{36}$$

以(35)及(30)代入(34)，得到

$$\frac{\mathrm{d}A_3}{\mathrm{d}z} - \frac{1}{2}(\alpha - \Gamma)A_3 = \left(\frac{C_0 \Gamma}{2A_1^*(0)}\right)\frac{\mathrm{e}^{\alpha z/2}}{1 + r\mathrm{e}^{2\alpha z - \alpha L}}. \tag{37}$$

设 $A_3(L) = 0$(在 $z > L$ 空间中没有相位共轭波)，则(37)式可以用常数变易法解出：

$$A_3(z) = -\frac{C_0}{A_1^*(0)}\mathrm{e}^{z(\alpha - \Gamma)/2} J(z), \tag{38}$$

式中

$$J(z) = \int_{\Gamma z/2}^{\Gamma L/2} \frac{\mathrm{e}^{\zeta}\mathrm{d}\zeta}{1 + r\mathrm{e}^{4\alpha\zeta/\Gamma - \alpha L}}. \tag{39}$$

以 $A_1 \sim A_4$ 在 $z = 0$ 的值代入(33)式，得到

$$-C_0 J(0) - A_2(L)A_4^*(0)\mathrm{e}^{-\alpha L/2} = C_0,$$

$$C_0 = -\frac{A_2(L)A_4^*(0)\mathrm{e}^{-\alpha L/2}}{J(0) + 1}. \tag{40}$$

代入(38)，得到

$$A_3(z) = \frac{A_2(L)A_4^*(0)\mathrm{e}^{\alpha(z-L)/2 - \Gamma z/2}}{A_1^*(0)}\left[\frac{J(z)}{J(0) + 1}\right], \tag{41}$$

以 $A_1 \sim A_3$ 的表达式代入(33)式，并用(40)式的结果，得到

$$A_4^*(z) = A_4^*(0)\mathrm{e}^{-z(\alpha + \Gamma)/2}\left[\frac{J(z) + \mathrm{e}^{\Gamma z/2}}{J(0) + 1}\right]. \tag{42}$$